DOCUMENT RESUME

ED 440 577 HE 032 793

AUTHOR	Massy, William F.
TITLE	Energizing Quality Work: Higher Education Quality Evaluation in Sweden and Denmark. Project 6, Quality and Productivity in Higher Education.
INSTITUTION	National Center for Postsecondary Improvement, Stanford, CA.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	NCPI-TR-6-06
PUB DATE	1999-00-00
NOTE	83p.
CONTRACT	R309A60001
AVAILABLE FROM	National Center for Postsecondary Improvement, Stanford University, School of Education, 520 Galvez Mall, 508 CERAS, Stanford, CA 94305-3084 ($7). Tel: 650-723-7724.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Educational Assessment; *Educational Quality; Evaluation Methods; Foreign Countries; *Higher Education; *Institutional Evaluation; *Quality Control; Site Analysis
IDENTIFIERS	*Denmark; *Sweden

ABSTRACT

This report, which is part of a study that seeks to help colleges, universities, and accreditation agencies improve their performance, presents insights gained from interviews and from documents collected in advance of and during visits to Denmark and Sweden in 1999 to examine higher education quality evaluations. Following Section 1, "Introduction," Section 2, "Antecedents and Approaches," discusses the two countries' approaches to evaluation (where they started and why they chose different strategies). Section 3, "Methodologies," focuses on the differences and similarities between the methodologies used in the two countries, and discusses preparing for evaluation, self-evaluations, user surveys, site visits, and reports and followup. Section 4, "Reactions to Evaluation," describes the institutions' reactions to the evaluation process. Section 5, "Impacts," assesses the impact of the two evaluation programs on quality work, discusses how evaluation affects academic autonomy, and comments upon the diffusion of quality work across institutions and their departments. Three monographs are appended: "Educational Quality Work (NCPI)" by William F. Massy and Andrea K. Wilger (1999); "Educational Quality Work in Denmark and Sweden"; and "A System for Quality Assurance and Quality Development" by Bengt-Ove Bostrom. (Includes 79 endnotes.) (SM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 440 577

Energizing Quality Work: Higher Education Quality Evaluation in Sweden and Denmark

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

National Center for Postsecondary Improvement

2

Stanford University • School of Education • 508 Cubberley Mail Code 3084 • Stanford, CA 94305-3084

HE 032 773

Energizing Quality Work: Higher Education Quality Evaluation in Sweden and Denmark

WILLIAM F. MASSY

National Center for Postsecondary Improvement
Stanford University
School of Education
520 Galvez Mall, 508 CERAS
Stanford, CA 94305-3084

The work reported herein was supported in part by the Educational Research and Development Center program, agreement number R309A60001, CFDA 84.309A, as administered by the Office of Educational Research and Improvement (OERI), U.S. Department of Education. The findings and opinions expressed in the report do not reflect the position or policies of OERI or the U.S. Department of Education. NCPI Technical Report Number 6-06.

This report represents work in progress. Please send comments to massy@stanford.edu.
Copyright © 1999 by William F. Massy.

Preface

NCPI Project 6, "Quality and Productivity in Higher Education," seeks to help colleges, universities, and accreditation and oversight agencies improve their performance. Despite the U.S. higher education system's manifest strengths and enviable reputation, my colleague Andrea Wilger and I believe it is time to ask hard questions about quality and productivity. Are U.S. colleges and universities doing the right things and are they doing them right? What changes would boost quality and access while containing cost? How can traditional institutions assimilate the new teaching and learning technologies and remain viable in today's intensively competitive environment?

Three forces are converging on traditional colleges and universities. Mission creep, the steady pressure to increase research at the expense of teaching, is driving costs up and reducing the delivered quality of education at all kinds of four-year institutions. Business's growing understanding of what it takes to assure and continuously improve quality is challenging higher education's traditional academic culture. Information technology is transforming on-campus instruction, enabling distance learning and new competition, and driving up costs. The confluence of these three forces is forcing the biggest transformation of postsecondary education since massification and the rise of research following World War II.

Our research has identified "educational quality work" as the key missing link in most institutions' transformational efforts. Quality work, defined in Appendix A of this report, refers to the activities of faculty, academic leaders, and oversight bodies that are aimed at improving and assuring quality. Quality work supports the provision of educational quality, but the two represent distinct kinds of activities. Quality provision depends on curricula and teaching and learning processes. Quality work focuses on the *improvement* of curricula and teaching and learning processes, and on *assuring* the best performance possible given available resources. We have concluded that improving educational quality work should be the single highest priority for institutions and agencies that wish to boost the quality of undergraduate education without spending more or decimating research and scholarship. We will in due course present the evidence for this conclusion and offer an action agenda in a book entitled, "Honoring the Trust: Quality, Productivity, and the Academy."

My visit to Scandinavia in June 1999 proved to be a transforming experience. The Swedes invented the term "quality work," which turned out to describe perfectly the concepts on which Wilger and I have been working for these past several years. The Swedish and Danish efforts to embed quality work into their university systems are highly relevant for U.S. institutions, as are the exemplary practices I observed and the

specifics of their evaluation methods. Sweden and Denmark have made real progress, and although much remains to be done, their experience can provide important lessons for U.S. readers and other readers around the world.

This report is long because I wanted to record my observations in considerable detail. Readers with limited time might best profit from reading Sections 1, 2, and 5, and Appendix A.

William F. Massy

Section One: Introduction

The worldwide trend toward institutional autonomy in higher education has become well entrenched during the last decade. State control has been transferred from input-oriented regulation toward criteria based on output quantity and quality. The trend is particularly noteworthy in Western Europe, where most tertiary systems were highly centralized until the 1980s. It also can be discerned in locales as diverse as Taiwan, Eastern Europe, and American state capitals. Tying resource allocation to student numbers or degrees represents a common first step toward decentralization, but governments soon recognize that this is not sufficient. Institutions should not only meet enrollment targets, they should provide the best education possible given their budgets. Hence higher education agencies are working overtime to establish or refine their quality oversight systems.

The Danish Centre for Quality Assurance and Evaluation of Higher Education and the French Comité National d'Evaluation report that in 1998 eleven European Union countries had established systematic national evaluation procedures and three others were in the process of doing so. Only the French-speaking community of Belgium, Greece, and Luxembourg seem not to have introduced such procedures.[1] Hong Kong, Australia, and New Zealand have operated formal evaluation systems for some years, and U.S. accreditation represents a type of evaluation. The International Network of Quality Assurance Agencies in Higher Education (INQAAHE), founded in 1993, attracted 220 participants from some 40 countries and regions to its Fourth International Conference in Santiago de Chile this past spring.

Denmark and Sweden established national quality evaluation systems in 1992 and 1993, respectively, and both have now completed a full cycle of evaluations. Unlike the UK with its polytechnic universities, neither country had much background in higher education quality assurance. They could think the matter through and start with a clean slate, aided by the growing body of literature and expertise that was becoming available during the early 1990s. Participants carefully documented their philosophy and experience, which makes the two countries an especially fruitful venue for research.

The two countries adopted different approaches. Sweden chose "academic audit" and Denmark chose "assessment." To summarize briefly, audit evaluates an institution's quality assurance and improvement processes, whereas assessment seeks to evaluate the actual quality of teaching and learning. Audit generally is conducted on an institution-wide basis, whereas assessment usually takes place at the level of a subject or discipline. Audit looks at "quality work" (the term used in Sweden to describe systematic and communicable approaches to quality assurance and improvement), whereas assessment seeks to evaluate quality itself.

Either approach can focus on improvement, accountability, or a combination of the two. Accountability provided the initial impetus for evaluation in both Denmark and Sweden, but both countries ended up stressing improvement. Energizing quality work turned out to be a key objective in the Danish assessments as well as in the Swedish audits.

Quality work and "quality oversight" have emerged as important elements of my research with Andrea Wilger on higher education quality and productivity. Quality oversight is our term for the evaluation of quality by an entity outside the university's internal governance apparatus. Appendix A presents our current thinking about quality work and quality oversight, and how they rationalize both internal and external evaluation and the goals of improvement and accountability. Taken together with teaching and learning processes, quality work and quality oversight constitute what we call the higher education "quality system."

The aforementioned factors led me to conduct benchmarking visits to Denmark and Sweden during June 1999. I visited the Danish Centre for Quality Assurance and Evaluation of Higher Education, the Copenhagen Business School, and the Danish Technical University. In Sweden I visited Göteborg University, Uppsala University, Stockholm University, the Parliament, and the National Agency for Higher Education. I talked with the evaluation programs' main architects and with the people responsible for quality work at the universities, but time did not permit meetings with a cross section of deans, department chairs, or faculty.[2]

This report discusses insights gained from my interviews and from documents collected in advance of and during the visits. It aims to inform U.S. and international audiences about the Danish and Swedish programs and their relevance to the work Wilger and I are doing at NCPI. Following this introduction, Section 2 discusses the two countries' approaches to evaluation—where they started and why they chose different strategies. Section 3 describes the two methodologies and their differences and similarities. Section 4 describes the universities' reactions to the evaluation process. Section 5 presents my own views on the methodologies and their impact on institutional behavior. Although I do offer evaluative comments, my sample was not sufficiently broad or deep to enable an independent analysis of impacts.

Section Two: Antecedents and Approaches

Sweden and Denmark started from similar positions and were motivated by the same kinds of forces. Staffan Wahlén, Coordinator of Quality Audit Programs at Sweden's National Agency for Higher Education, describes these forces as "...a change of rela-

tionship between the state and the universities, a move from management by rule to management by goals, objectives, and results."[3] He goes on to list three particularly important elements of change.

1. The universities are developing far more *autonomy*. In Sweden, the size of the higher education ordinance is now one-fourth of what it was six years ago. Universities are now more or less entirely responsible for programmes and courses (within the general framework of the degree system), management at all levels, recruitment, appointments, etc.

2. There is an emphasis on the *professional* role of university staff. In an autonomous, or quasi-autonomous system, with few rules imposed by external forces, there is a need for a professional culture to develop in which professional groups (i.e., professors, lecturers, other staff, university leaders) take full responsibility for their work and its results.

3. *Self-regulation* becomes important for continual internal development. Universities are expected continuously to follow up and evaluate their own activities and take action on the basis of the results.[4]

Similar forces were operating in Denmark. How these common antecedents evolved to produce different strategies illustrates some important lessons of evaluation.

Sweden

Sweden's 39 institutions of higher education enroll almost 300,000 students, up from only about 37,000 in 1960, and serve a population of about 9 million. The system consists of traditional universities, colleges offering basic degrees, and colleges of fine arts. Uppsala, northern Europe's oldest university, was founded in 1477, whereas a number of other institutions achieved university status only recently. A bachelor's degree requires three years of study, though it often takes longer. Masters degrees are planned for one year and the doctorate is planned for four years. The state provides most of the funding. Systems of program review and accreditation exist but these do not intersect with the audit system.

Faculty viewpoints appear fairly traditional, with a strong emphasis on research. Criteria for scholarship and approaches to teaching and learning are broadly similar to those in America. Until 1992 the dominant representation of students and faculty on governing boards complicated decision making, but this was a moot issue as long as most matters were decided in Stockholm. The content of degree programs was traditionally determined nationally, and political ideology holds that all institutions have the same

quality. Considerable institutional diversity can be observed, however. Students try to get into the "hottest" institutions, where hottest is defined more in terms of prestige and location than educational quality. The degree of competition is described as "fierce," with only 40 to 50 percent of students getting their first choice.

Higher education decentralization began during the 1980s as part of a national move toward local governance—a trend aimed at boosting responsibility and efficiency. It accelerated sharply in the 1990s when rectors and governing boards were given more authority and resource allocation and quality assurance were restructured. (Until 1993, for example, each new and continuing professorship had to be approved by the Education Ministry.) Boosting responsibility and efficiency became particularly important as economic recession and reaction to a lagging participation rate forced institutions to accommodate doubled course enrollments and 40 percent more students with only 20 percent more staff.

Questions about educational quality, and the system's initial reactions to them, also fueled demands for change. For example, one respondent described many humanities and social science faculties as being "...almost in crisis...with very few lessons per week—mostly large lecture classes with few tutorials. Both students and faculty were very stressed during lectures." These problems did not go unnoticed:

> In 1989, the Social Democratic Minister of Education responded to criticism from students regarding poor teaching and called upon a commission to scrutinize the state of the art in teaching in undergraduate education and to suggest measures for its improvement. Their point of departure was not political visions, but was to identify shortcomings.[5]

The Commission's 1992 report, "Freedom, Responsibility, Competence," helped open the quality dialog but it was rather conservative in its recommendations. Indeed, "Many academics were almost flattered by the many references to Humboltian ideals, and to the declared need for a connection between teaching and research as major criteria for increased quality in higher education."[6] The dominant viewpoint within academe seemed to hold that an "invisible hand" would produce educational quality as long as research quality could be maintained. But in 1993 the traditional academic view would be challenged by the incoming Conservative Party's Minister of Education, Per Unckel.

The new Minister and his constituents viewed Swedish higher education as a system charged with producing value for money rather than as a collection of individual institutions that should be sustained and protected for their own sake.[7] His first reform was to change the resource allocation system "to pay off according to what you get." In broad terms, the Minister now sets funding according to the number of FTE students

and the number of credits earned by the students. (Institutions negotiate target numbers with the Ministry. Enrollment can exceed the targets but institutions don't get paid for the excess. Failure to meet the targets results in funding penalties.) This process was designed to take the quality of teaching into account, since poor instruction presumably reduces the rate at which credits are earned. Skeptics point to confounding with incoming student quality and the possibility—considered explicitly by government but judged remote by most academics—that standards might be reduced to improve funding. Despite these concerns, the system has proven durable and for the most part has been well accepted.

The other reform agendum targeted educational quality directly. Staffan Wahlén describes the situation this way.

> It was thus made clear that each institution was responsible for maintaining and improving the quality of its own activities, and was accountable to the Government and society for this. It may be maintained that universities and colleges have always been quality-driven. What has now been added, however, is that they must have (and demonstrate that they have) systematic improvement processes regarding undergraduate education, graduate education, research, and administration. They are required to develop routines for reflecting on their activities, and make corrections wherever necessary for the sound improvement of the institution."[8]

Legislation called for the development of quantitative performance indicators to "score" the quality of education. "Inferior" scores would deduct five percent of the institution's funding. (The original proposal called for eight to ten percent, but Parliament reduced this to five percent.) "Acceptable" scores would deduct 2.5 percent, while "perfect" scores would lead to full funding. Thus the institutions were to be held strictly accountable for quality with enforcement via the funding process.

In 1993 an Office for Evaluation of Higher Education, soon known as the Office of the University Chancellor, was established. Stig Hagström, a Physics Professor at Stanford University and former Xerox executive, was recruited as Chancellor.[9] In 1995 the Office merged with three other agencies to form the National Agency for Higher Education. Its tasks include audit and assessment, but also (for example) legal supervision, information to students from a national perspective, and certain supportive functions such as a Council for the Renewal of Undergraduate Education and a program for academic leadership development. Hagström was succeeded on January 1, 1999, by Sigbrit Franke, Rector of Umeå University and a former Professor of Education.

The Chancellor quickly concluded that no system of quantitative performance indicators could be valid for all institutions and subjects. Moreover, the universities would not participate wholeheartedly in subjective quality review as long as the direct linkage to funding remained in place. They would fight to avoid any deduction, so the self-disclosure of subtle quality issues could hardly be expected. Hagström and his colleagues became convinced that effective faculty buy-in was essential, and that starting the quality review process on an adversarial basis would make future progress doubly difficult. ("If I show the stick up front, I'll never get the process started."[10]) It also became clear that because the evaluators and academics would be learning on the job, the process, the criteria, and the quality of institutional responses would evolve during the first evaluation cycle. This would complicate the comparison process and undermine the case for any particular financial penalty. Hagström convinced Unckel to drop the direct linkage to funding, at least for the initial phase of the project. The latter's good relations with the Finance Ministry carried the day, though at the cost of some political capital. Skeptics felt that "the institutions just didn't want to be compared," and that "Hagström wasn't tough enough on quality."

The Chancellor's second decision committed the Swedish quality review process to academic audit rather than assessment. The audits would focus on undergraduate education and cover the following main areas:

- *The strategies of the quality enhancement process:* What policy, plans, and programmes have been developed to realize the overall goals and ambitions for quality enhancement? Have the goals been operationalised? What form has been given to the organisation and distribution of responsibility? How have priorities been set?

- *Leadership:* How is leadership exercised on different levels to, for example, impart visions, create motivation, participation and responsibility, develop competence, and strategically implement and follow up on quality enhancements programmes?

- *Co-operation with stakeholders:* In what ways have the stakeholders been identified, their needs and demands determined, and how has the institution co-operated with them?

- *Involvement in quality enhancement processes:* How and to what extent are teachers, researchers, administrative staff, and students committed, involved, and responsible participants?

- *Integration:* How is quality enhancement integrated into university work and its various components?

- *Systems of evaluation and follow-up:* What methods, routines and measures have been adopted for recurring evaluations and the resultant development?

- *External professional relations:* In what ways is the university pursuing national and international contacts of long-term and permanent importance for the professional nature and future direction of university activities?[11]

The seven points illustrate how an audit of quality work focuses on institutional and departmental processes for assuring and improving quality, rather than on the level of quality itself. The questions apply to any kind of institution, and many of them also can be asked of academic subunits such as departments. To assess quality, on the other hand, one must test educational processes and outcomes against standards appropriate to the discipline and the institution's mission.

Several considerations favored Sweden's choice of audit over assessment. To paraphrase Hagström's account,

> What happens if you decide after assessment that something needs changing? You're at a loss without knowledge of how the institution assures and improves quality. Therefore, it made sense to start by looking at quality work. We felt that assessments might be done later, but a subsequent change of government made that impractical. [12]

A second factor favoring audit was that it is less threatening than subject-level assessments. Audits address systematic issues whereas, in assessment, departments or even individual professors can be accused of poor teaching. Third, by announcing an audit of quality work the Agency could connect initially with the small number of informed and committed quality enthusiasts who comprised the "quality units" at most institutions. These enthusiasts could work from within to establish the quality principles in as many departments as possible, prior to the audit. Addressing educational quality directly through assessment might have focused attention on what the departments were doing already, rather than amplifying the influence of the central quality units. Finally, the Chancellor wanted to make the first-round evaluations strictly improvement-oriented—in contrast the ongoing UK experiment which had associated assessment with accountability in the minds of many. [13] Shifting to an audit focus helped alleviate fears that the Swedish evaluations would, after all, be summative. Respondent after respondent said the program's "light touch" and thoughtful demeanor made it easier to engage professors in a real dialog about quality work.

The Swedish audit process bears some resemblance to the one developed by the UK's Higher Education Quality Council (HEQC), but there are important differences. For example, the Swedish program places less emphasis on formal quality assurance procedures. It does not appear that the Swedish design was informed to a significant degree by the HEQC experience. Indeed, Hagström indicates that the basic design was motivated more by his experience with quality processes at Xerox than any precedents from higher education. (Agency staff members were well informed about current quality trends in higher education, however.[14]) The Swedish process has much in common with Hong Kong's successful quality process audits,[15] although the two evolved independently.

Denmark

Denmark has a population of about 5.3 million and enrolls about 160,000 students. There are 12 universities, of which 5 are traditional multi-faculty institutions and 7 are mono-faculty (technical university, business school, etc.). Denmark follows the British model of a three-year first degree. Masters degrees are planned for two years and the doctorate for three years. In addition, sub-university institutions provide a large number of "medium level" three-year programs (nurses training, social workers training, etc.), with student numbers approximately equal to those at the universities. Many Danish programs appear more specialized than their Swedish counterparts. Five powerful National Education Councils (for humanities, social science, engineering, science, and health science) advise the Minister of Education on program content and other disciplinary matters, and also on general higher education policy issues. The programs are called *uddannelse*, which translates roughly into "subject area" or "subdiscipline."[16]

Reform of Danish higher education began in 1992 when the conservative-liberal minority government achieved a number of far-reaching compromises with Parliament. The principles of reform stressed freedom and autonomy with accountability—that is, "deregulation and decentralization, combined with mechanisms to ensure quality."[17] In addition to calling for stiffer admission requirements to the most demanding programs and additional examinations along the road to a degree, the reform strengthened quality assurance by creating an independent Centre for Quality Assurance and Evaluation of Higher Education and reorganizing the system of external examiners. The Centre, which began operating in July of 1992, is an independent agency funded by the Ministry of Education and governed by the five chairpersons of the national education councils. Staffing consists of a director, a deputy director, eleven academics, and fifteen full- and part-time support people. Professor Christian Thune, former Dean of the Faculty of Social Sciences at Copenhagen University and Chairman of the National Education Council for Social Services, was appointed as its first director.[18]

The Centre's mandate, which covered both university and medium-level programs, was as follows.

- to initiate evaluation processes for higher education in Denmark;
- to develop appropriate methods for *assessing programs;*
- to inspire and guide the higher education institutions in aspects concerning evaluation and quality; and
- to compile national and international experience on evaluation of the educational system and quality development. [19]

Denmark's choice of assessment as its evaluation method seems to have been determined very early in the higher education reform process—probably even before the Evaluation Centre began operations. The mandate's second element speaks of "assessing programs," not auditing quality work. In 1992, most discussions of evaluation focused on assessment. (The British Higher Education Quality Council's audit program was just getting off the ground and the audit programs in Sweden and Hong Kong would not be launched for another few years.) Given the sense of the times, it was natural to address the quality of teaching itself rather than the meta-concept of quality work.

As evidenced by the composition of the Centre's governing board, the five national education councils were viewed from the beginning as its primary clients. This would have impeded any decision to focus on quality work rather than subject-level assessment. The Councils were interested in the cross-institutional comparisons across a whole *uddannelse* that can be obtained from assessment. Because quality work at the subject level cannot really be separated from that of the institution, to concentrate on such work would have led the Centre toward institution-level evaluations—which lie outside its mandate. For these understandable reasons, the question of whether to do audits instead of assessments never seems to have arisen. However, the Centre's director confirms that the process did come to include a "very distinct audit focus."

According to Thune, the Centre's goals included "raising the priority of teaching and learning, and increasing the leverage of people inside the institutions who care about them." He cites problems with the Danish external examiner system, student passivity, and most of all the faculty's preoccupation with research. ("The new generation of faculty cares about teaching but it still takes research to get their ticket punched.") Indicators of success would include "more explicit, systematic, formal quality mechanisms"—i.e., "quality work" as defined herein.

Further insights into the Centre's goals can be obtained from Thune's commentary, circa 1995, about the institutional benefits expected from evaluation in the new climate of autonomy.

> The presidents, deans, and governing boards were now facing independent, broad, and often difficult decision-making. Systematic evaluations would provide the institutions with an insight into the quality of their own study programmes. Good evaluations, which reflected the relation between institutional goals and realities, could therefore form the basis for planning and priorities of tasks. [20]

By covering all the requirements for educational quality—including inputs, program structures, teaching and learning processes, and outcomes measures—the Danish-style assessments were to provide broad-spectrum guidance to both government and institutions.

Thune insisted that improvement and accountability be combined in the design of the evaluation program. He describes the theoretical difference as follows.

> The criteria for *improvement* are those procedures that are conducive to strengthening the conditions, motivations, scope and level of information of the higher education institutions in this direction: in other words, procedures that engage the institutions in a self-learning process. Such procedures should aim at promoting future performance (formative evaluation) rather than judgments about past performance (summative evaluation). They should lead to ends that are specifically in the interest of the institutions and towards the specification of quality according to criteria that are internal or may be made internal by them.
>
> The criteria for *accountability* are procedures that lead to the assessment of the quality of teaching and learning in terms of criteria set down by external authorities and institutions and with the goal of strengthening external insight and even control, opening the door for eventual corrective action. [21]

Mixing improvement and accountability ran counter to most expert opinion of the time, which held that government-owned programs tend toward bureaucracy and emphasize accountability at the expense of improvement. On the other hand, most public officials felt they could not rely solely on university-owned programs to provide the depth and objectivity needed to discharge government's obligations to its constituents.

The new Danish Centre was publicly owned and therefore sensitive to accountability goals, but it also wished to respond to the institutions' need for improvement. Thune and his colleagues felt that improvement and accountability could, to use his words, be "merged or

synthesised in a dual approach with an emphasis on improvement."[22] The fact that the government reform package did not directly link evaluation to funding, plus careful thought and wide consultation by the Evaluation Centre, made their strategy feasible.

Section Three: Methodologies

While the approach to evaluation differs, the basic conception of quality and the core elements of the two countries' evaluation processes are surprisingly similar. Sweden describes quality as a "dynamic force for change" rather than conformance to a "static concept... defined by predetermined quality parameters... Quality is neither an objectively determinable nor an unchanging property of a phenomenon. Quality is a judgment about this phenomenon, made by a stakeholder or interested party—on a given occasion—grounded in a subjective judgement of its value to him or her."[23] Denmark reflects these sentiments in its mistrust of quantitative performance indicators, emphasis on stakeholder inputs, and focus on processes for assessing and improving quality within each *uddannelse*.[24]

Both processes include preparation for evaluation, an institutional self-study, a site visit, and a public report and follow-up. Denmark also includes a stakeholder survey. Each Danish evaluation is led by a "steering committee," usually with five members, which also serves as the site evaluation team. The Evaluation Centre provides staff support to the steering committee. In Sweden the National Agency manages the process, with the five-member audit teams being responsible for the site visit and final report.[25]

Preparation for Evaluation

In Denmark, evaluation means comparing all the institutions that teach a particular *uddannelse* (typically a half-dozen or so). The report goes to the National Education Council responsible for that discipline. Each exercise begins with a written contract, called a "*kommissorium*," which particularizes the Centre's general guidelines by describing the terms of reference for the evaluation. The terms include the specific subfields to be evaluated, criteria to be used, and division of responsibility among the parties, logistics, timeframe, and any special issues that will be raised in the evaluation. The relevant education council is invited to propose issues, and the Council must approve the *kommissorium* before the evaluation can proceed. (In practice, the *kommissorium's* content does not vary greatly from evaluation to evaluation.) The *kommissorium* also was described by Centre staff as "a way of controlling the experts." That is, it provides the steering committee with boundary conditions and a checklist of items to cover without constraining its ability to reach independent conclusions.

The Danish Centre maintains a "Quality File" for use in program planning. The file contains detailed descriptions of every element of the evaluation process. These descriptions are described as very well organized and complete, and the file is updated regularly. The resulting high level of documentation aids the Centre in preparing strong *kommissoria* and helps make it an effective learning organization.

In Sweden, where an evaluation addresses a single institution, the terms of reference are contained in the program's general methodological documentation and in the memories of participants in a planning meeting between the institution and Agency staff. The Swedish Agency maintains some checklists and templates, but as of early 1998 it had no comprehensive quality file. While such a file was viewed as desirable on general grounds, the simpler stakeholder structure and more circumscribed site visits reduced the need to some extent.[26]

Both agencies arrange orientation meetings with the institutions being evaluated. These sessions typically involve academic leaders and the people responsible for the institution's quality program and for preparing the self-assessment. In Sweden, the orientation meeting involves university administrators, the chair and secretary of the audit team, and the Agency liaison person assigned to the audit. The Danish meetings are similar except that they do not include anyone from the site visit team. The meetings discuss the evaluation's terms of reference, address logistical issues, and discuss expectations for the self-assessment. They provide important venues for dispelling misconceptions about the nature and purpose of the evaluations as well as for dealing with nut-and-bolts issues.

The Danish evaluations include a formal request for materials, to be supplied by the institutions prior to preparing their self-evaluations, about the structure of the *uddannelse* and the relevant teaching programs. An Agency staff member analyzes these materials and prepares a briefing document for review by the steering committee at its first meeting. The document provides committee members with a common base of understanding about the programs to be evaluated. While some felt the staff's subjective interpretation might open the door to bias, the committee members' expertise appears to have mitigated any such effects.

Self-Evaluations

Both Sweden and Denmark consider self-assessment to be an essential element of evaluation. As in virtually all evaluation systems, it plays a dual role: as an instrument of internal quality improvement and as the most important document for the site visit team. Both countries emphasize the formative aspects. "The more self-assessment is given priority in the process, the more it will function as training and preparation [to help] the institution

or study programme take over responsibility for its own quality development—and the less it will be seen as producing information for the expert committee."[27] Both countries stress that the self-assessments should be "not only descriptive, but truly analytical."[28] Among the desirable attributes mentioned are providing a framework for defining and assessing quality, balancing qualitative and quantitative data, and describing the program's strengths, weaknesses, and strategies for improvement.

Neither the Evaluation Centre nor the National Agency specified the process by which the self-evaluation reports were to be prepared. However, Denmark "strongly advised but could not dictate" that the reports be prepared by a group including leadership, staff, students, and administration. They wanted to secure maximum co-ownership, and based on follow-up reports requested in the case of some evaluations they believe that a majority of institutions did follow the advice.

Sweden's National Agency views the format of the self-evaluation, and to a large extent its content, as the responsibility of the institutions themselves. Universities approach quality differently, and each institution should "find the forms appropriate for its own evaluation."[29] The Agency's *Guidelines for Institutions* states simply that the self-evaluation should not exceed 25 pages and that it should contain:

- a presentation of how the self-evaluation was conducted;
- a description (based on the quality enhancement program) of the institution's quality program, indicating major goals, quality targets and preconditions, strategies and prioritized areas, together with the structural components of the quality program, their implementation and results;
- an analysis and evaluation of the university's own quality program and activities, as regards their strengths and weakness, and their threats and opportunities; and
- conclusions of the self-evaluation, as well as the measures the institution intends to take to enhance its quality program and improve its results.[30]

Institutions are invited to a one-day workshop aimed at amplifying the audit guidelines. In addition to staff presentations, representatives from institutions that have already carried out self-evaluation are invited to share their experiences and convey advice. Topics discussed include:

- What do the terms "quality work," "professional reflection," and "critical self-analysis" signify? What should a self-evaluation contain?
- How can a self-evaluation be planned and carried out? How do you inform and engage staff and students?
- What problems and possibilities can turn up during the work of the self-evaluation?
- How does a self-evaluation of quality work differ from a self-evaluation of quality as such? [31]

The audit team participates in a further seminar with the institutions after the self-evaluation has been substantially completed but before it is set in final form. Both seminars were described as helpful in clarifying the purpose of the evaluation and filling in gaps of coverage.

Stockholm University, which was evaluated late in the cycle, asked departments to formulate goals for three years and then report on the progress during the first year and how the goals were being revised for the next rolling three-year period. Not all self-evaluations were this well focused, but Agency staff members reiterated that they did not want to be overly prescriptive with respect to the self-evaluations.

Denmark took the opposite approach by providing very detailed specifications for the self-evaluations. The Centre prepares a "Guidance to Self-Evaluation" document for each assessment, based partly on general policies contained in its operating manual and partly on subject-related issues. The specification includes quantitative performance indicators and 25 qualitative categories described as containing 2-8 questions each. (It seems similar to the typical American accreditation manual in terms of the amount of information sought.) The specification needs to be followed precisely. According to one commentator, "The Centre's name for 'guidance' is to some extent misleading since all the questions should be answered, in the right order." [32]

While the Danish specifications are constraining, they provide some significant benefits as well. For example, a Centre respondent said that a good report will "leave very little room for maneuver," and that the teams "do a text analysis to determine whether the report is sufficient." I did not encounter criticism of the self-evaluation design specification at the two institutions I visited. According to one respondent, "The 60-plus questions forced us to think things through. The Self-evaluation was the most important part of the assessment process."

User Survey

Denmark's Evaluation Centre has taken the unusual step of conducting its own user surveys. Four groups of users, also known as stakeholders, are of interest: employers, former students, current students, and the so-called external examiners. (The external examiners—professors from other universities and representatives from industry, government, high schools, etc.—help the institution set appropriate academic standards and directions.) The Centre chooses one of these groups for each evaluation, and then commissions a consulting firm to design and implement a mail or focus group interview survey. It appears that employers are surveyed most often and the external examiners least. The steering committee can obtain information directly from current students and external examiners but, as described later, the sampling of employers and former students can be problematic.

Employers are asked what characteristics of graduates are most important. In addition to specific professional qualifications, relevant characteristics include the integration of specific and general knowledge, communication skills, and problem solving. Additional questions probe whether the students are getting better or worse, and whether any institutions have consulted the respondent's firm about program criteria and performance. The questions cover the *uddannelse* in general, not the programs of the individual institutions.

Site Visits

As in all evaluation systems, the site visit allows team members to check the institutions' documentation and form their own impressions on the ground. Sweden varies the length of the visit depending on the institution's size and complexity, but two to three days represents the norm. Denmark has standardized one day per institution. Recall, however, that the Danish evaluations focus on a single *uddannelse* and may involve as many as ten institutions.

Swedish audit teams typically consist of five members plus the liaison officer from the National Agency. Positions represented include professors, heads and ex-heads of institutions, representatives from industry, students, and a university administrator (who usually serve as the team's secretary). Most auditors are drawn from Swedish universities, though foreigners may be used when language difficulties can be overcome. Liberal use of university presidents and top administrators provides a sense of authority and acceptability. Team members are appointed by the Agency after consultation with the institution, and the teams are activated just before the site visit.

The Danish steering committees (which also serve as the site visit teams) always include significant representation from trade or industry. They are likely to include foreign academics with domestic academics appearing less frequently. In the words of the Director, "The priority was on independence and integrity so we tried to minimize the number of Danish university staff members invited to join and to bet on Norwegians and Swedes (since 1994 at least one Nordic member was included in each panel) and employers' representatives." This priority was particularly important given the small size of the Danish system and the fact that, starting up to six months before the site visit, the steering committee takes full responsibility for the subject-matter side of evaluation. The *kommissorium* states the division of labor clearly: "the experts have the responsibility of the subject specialist and the Centre's staff the responsibility for the process methodology."

Successful site visits require careful preparation. The Swedish *Guidelines* describe a two-step process. As a first step, members of the audit team work independently to evaluate the submitted materials in terms of the following kinds of questions:

- Are the university's goals and strategy for the quality program clearly formulated?
- Is the self-evaluation sufficiently analytical, e.g., are strengths and weaknesses clearly formulated, examined, and documented?
- Does the paper contain a clear description of the measures the institution intends to take to deal with its weaknesses?
- Is there any information missing in the material submitted by the institution?

In the second stage, the group works together to achieve a consensus about more detailed questions like these:

- Is there agreement between the university's goals, underlying ideas and ambitions, the operational frame, and the strategic approach?
- Do the strategy and its guidelines, in the form of its quality policy, goal specifications, priorities, plans, organization and resources, agree with the practical implementation of the quality program?
- Does the quality program contain the requisite structural components (in the form of quality routines, evaluation systems, and the like) that will enable the institution to attain the specified goal?

- Do the quality program and its implementation work well as regards factors such as leadership, participation, and cooperation with external stakeholders?

- Do the operations achieve concrete and goal-related results in the form of improvements that have changed quality levels in undergraduate education, research, graduate education, and other tasks of the institution?

- Does the quality program encourage consideration of and adjustment to the needs and demands of external organizations, groups and authorities in a relevant manner? [33]

Discussion of these questions facilitates preparation of a detailed agenda for the site visit.

Denmark follows an equally rigorous procedure. For example, the visit team asks both the permanent and temporary faculty whether the self-evaluation reflects reality, then goes on to probe their objectives and their linkage to teaching and learning. I did not try to obtain the protocols, which are in Danish and would vary by subject in any case. It appears that the protocols cover a broad range of issues, including questions, like those above, aimed at assessing quality work.

Both countries attach great importance to the training of panelists. Academics and stakeholders possess detailed knowledge of their disciplines and domains of experience, but this rarely extends to the process of evaluation. The ability to analyze a self-evaluation report and to ask probing and fair questions during a site visit is an acquired art. Both the Evaluation Centre and the National Agency for Higher Education expend considerable thought, time, and resources in the design and implementation of training regimens.

Reports and Follow-up

Both countries consider it essential that the evaluation reports be published. One reason is of course rooted in accountability: a privately held report cannot inform government or the market. Another reason can be found in the systems' improvement objective. Most quality improvement, and virtually all aspects of quality work, requires extensive dialog within the institution. A closely held report cannot stimulate and inform such dialog for the simple reason that not enough people will know the particulars. To distribute the report to sufficient numbers of faculty, administrators, and students to achieve its formative goals is tantamount to making it public. One might as well publish at the outside to avoid the possibility of losing the initiative through leaks to the press.

Achieving departmental acceptance and providing maximum guidance for improvement requires careful documentation of conclusions and recommendations. Substantial resources must be put into the report-writing process—more, probably, than can be afforded by the panel chair. In Denmark the Centre representative who participates in the site visits takes primary responsibility for preparing the first draft of the report.

Exit conferences at the conclusion of the site visit are common in most evaluation schemes, and Sweden includes such a conference. Denmark does not, largely as a matter of principle but also because the short duration of the Danish visit (one day per institution) would make inclusion difficult. The Centre believes that the site visit team will not have had enough time to reflect on its experience, that a preliminary verbal assessment may well be too positive, and that there can be no basis for comparison until all institutions involved with the *uddannelse* have been visited. The Centre's project leader begins work on each institution's section of the report shortly after the site visit, but the drafting remains in progress until the entire visit program has been completed. Then the project leader provides a draft to the steering committee, which typically meets two or three times before finalizing its preliminary report.

Denmark convenes a separate conference, well after the site visit, to discuss each evaluation's preliminary report. The conference consists of about four representatives from each evaluated institution, the steering committee, the project director, and the Centre Director. Institutional participants receive the preliminary report about two weeks in advance of the meeting. The meeting opens with a short presentation by the chair, after which each institution gets about fifteen minutes to comment. (The comments also are presented in writing.) All parties engage in general discussion, after which the steering committee meets in executive session to finalize the report. Some accounts indicate that the conference has turned out be less than fully successful—that it doesn't function very well because the competitive atmosphere makes it hard to get a real discussion going. However, the Centre wants the experts to confront the institutions and no better way has yet presented itself.

Sweden convenes no conference, though the draft report is sent to the institutions for correction of factual errors before publication. A post-report conference would be less relevant in the Swedish context because only one institution is evaluated at a time and an exit conference is provided at the end of the site visit.

The Danish reports conform to a fixed format for their introduction, executive summary, English summary, and included appendices. More flexible formats characterize the sections on analyses, conclusions, and recommendations—which usually address both institutional and national issues. The Centre provides free copies to the relevant National Education Council and to the institution's self-evaluation group. Additional

copies can be purchased for wider distribution. The sharpness of the analyses and the level of detail in the recommendations have been described as variable, and also shorter than their Swedish counterparts.[34] However, some think the Danish reports are more independent and hard-hitting.

After each Danish evaluation, the national council provides the Education Ministry with an official letter commenting on the report and citing actions it believes should be taken. Since July 1, 1997, each institution must provide the Ministry with a plan for following up on the evaluation's recommendations and the points in the Council's letter. In Sweden, the Agency (not the Ministry) requires a "one-year follow-up" from each institution audited. These documents report on what has been accomplished since the audit and then go on to discuss plans for the future. Agency staff members typically visit the institution to discuss the follow-up report.

Section Four: Reactions to Evaluation

Accounts of the Swedish and Danish experience describe the participants' reactions to evaluation in considerable detail, and I was able to supplement this material with my own interviews. Participants include people within the institutions, members of the evaluation teams, and representatives from the evaluation agencies. Their reactions are valuable for determining how the two countries' evaluation processes might be improved and for gaining insight about the consequences of evaluation.

Overall Reactions

Both Sweden and Denmark commissioned independent reviews of the impacts of their systems and their impacts. Bjørn Stensaker of the Norwegian Institute for Studies in Research and Higher Education (Oslo) reviewed the Swedish Agency, and an international expert panel that included Stensaker and Professor Berit Askling of Uppsala University reviewed the Danish system. The reviews included significant numbers of interviews—for example, with university presidents, deans, program supervisors, and department chairs. The Danish Minister of Education also asked a major consultancy firm to study the impact of evaluation on the institutions.

The overall reaction to evaluation appears to have been favorable in both countries. To quote the Danish experience,

> The experts concluded that the Centre had fulfilled the objectives outlined in the mandate in a thorough and systematic way, and very importantly, that the Centre had been successful in setting up evaluation procedures

> which combined the two objectives of control and improvement. The large majority of stakeholder's parties [sic] found that the evaluations carried out by the Centre were valuable, and that they had been carried out in a professional way and according to appropriate methods.[35]

Similar statements could be made about the Swedish system. The main criticism involves faculty time consumption, but it appears subdued relative to faculty opinion about most state assessment and regulatory initiatives in the United States.

My interviews and literature review also produced generally favorable reactions. While I focused on the evaluation agencies and people within the universities that are knowledgeable about and thus sympathetic to quality work, it is hard to believe that a groundswell of dissatisfaction would have escaped detection.

Swedish National Agency staff member Malin Östling, who spent three months at the Danish Evaluation Centre during the fall of 1997, summarizes the reaction as shown below. She interviewed 17 people at 8 Danish institutions that had been evaluated: two in law, three in machine engineering, and three in Danish/Nordic languages. (Italics indicate quotations from her respondents.)

> Some of those interviewed claimed they had a positive view of the evaluation before it began and that this has remained during and after the process. Most of them also believed that the evaluation would lead to new insights and concrete changes.
>
> *My hope was that we would get an analysis that we could use in future work. That is what happened, too.*
>
> Many interviewees were negative before the evaluation, however. The main reason seems to have been an anxiety that it would take too much time from other activities, which many think they have.
>
> *I was very negative before the evaluation. I thought it would take a lot of time and energy, and not give the equivalent in contribution of new ideas. Unfortunately, that was the case.*
>
> Then there were those who had a hard time accepting the time consumption but still can see that they got something out of the work. There are also expressions of anxiety about an external organization assessing the activities, and what the consequences would be if the *uddannelse* appeared to be of an unsatisfactory quality.

I realized that [the evaluation] would be a big job—and so it was. Then I wondered how deep they would go, how close to the truth they would get, if the steering committee would detect the deficiencies that we did not talk about openly in the self-evaluation. I hoped to get support to carry changes through that I think are important but that I have trouble getting support for here.

I thought it would take a lot of time—it certainly did. It was troublesome. But we were going to do it—there was no way out. I had read previous reports, and had seen that they [those evaluated] had a chance to review their work and thought that it surely would give something to us too.

None of those interviewed changed their opinions about the evaluation from positive to negative after it was completed.[36]

One of my Danish institutional respondents summed up the process's strong points as: "the self-evaluations, the avoidance of institutional rankings, and the fact that the evaluations were open to the public at every stage." The only criticism was that the reports should have taken the *uddannelse*'s economic circumstances into account when recommending change. (Denmark's reports reflect a self-denying ordinance that the teams will not comment on resource allocation. In particular, they will not recommend that additional funds be provided the *uddannelse*.)

The Swedish reactions appear generally similar. The audit was viewed as helpful despite the time required and a drumbeat of concern that the definitions of quality work and the guidelines for self-evaluation were ambiguous. In the words of one institutional respondent:

- Audit put quality work on the agenda—the Agency asked each institution to have a quality program and then audited it. Furthermore, the requests legitimized the central administration's quality work. [This was particularly important because the government's] resource allocation procedures tended to bypass the central administration and go directly to the departments.

- The way audit was conducted turned out to be confusing, however. The guidelines were very open. The Agency wanted the universities to be learning organizations, but more attention should have been paid to what that meant. Too much was left to the audit teams, which led to uncontrolled variation in criteria. (This was a very weak element.)

Other respondents echoed these sentiments.

On balance, it appears that the audits created value and were not too costly or intrusive. While some begrudge the time required, there is a broad consensus that on the whole the exercise was worthwhile though not utterly transforming. In the aforementioned review of the Swedish system, Bjørn Stensaker reports as follows.

> Summing up the answers from the interviews [with department heads], twenty-one of twenty-four respondents thought the audit was concretely *important* for the work of their departments, but that the changes, adaptations or measures that were implemented were often of an incremental nature. In the same way that effects of assessments appear to come about, the respondents stressed the importance of the whole *process*, where one sees the effects as "the result of anticipations, the preparation of the self-evaluation, and the interaction with the external auditing team (and with colleagues) as well as the audit report and follow-up actions."[37]

He goes on to report "great variations among departments" and that "the *greatest* effects were found [in the] management and organization of quality assurance work [and the] development of academic staff." I believe the same applies to the Danish program and that it represents a favorable "bottom line" for any evaluation exercise.

Self-Evaluations

As in most countries, reaction in Denmark and Sweden indicates that the self-evaluation is potentially the most useful element of the evaluation process. However, feedback in both countries indicates a certain spottiness of implementation. Östling's Danish interviews sum up the tension generated by the self-evaluations:

> Many people think [the self-evaluations] provide an opportunity to increase internal knowledge [about the *uddannelse* as taught]. Others see this purpose but think it has not been fulfilled.
>
> > *We really wanted to get something out of [the self-evaluation] for ourselves—since it was a big job. So we tried not to think that it was meant for the steering committee, because then we would not have been so open.*
> >
> > *I assume the thought is that we should "discover ourselves." But the statistics did not give us anything…, not in the way [the instructions] were drawn up.*

> *The implicit purpose is to turn on some processes at the university and make it easier to change things afterwards because they give a more obvious picture of the situation, at least for the faculty management. Of course we knew about our problems before, but we had not put words to them earlier—and that was a good thing to do.* [38]

The last quotation appeared a number of times. Respondents said that the problems were well known to the "inner circle," but that the self-evaluation brought them into the open and increased their saliency for departmental and decanal decision making.

Östling also conducted extensive interviews at Swedish institutions. On the subject of self-evaluation she reports the following positive responses.

> *We are totally satisfied with the self-evaluation report and even more so with the work on it. It was in some respects arduous but gave us, in the end, a considerable active interest in and awareness of the importance of quality work.*
>
> *We know that our self-evaluation report perhaps is not what the National Agency wanted, but it was and is useful to us, which is important, isn't it?*
>
> *Most of us are content with the self-evaluation report even though it is a document of compromise.* [39]

The reactions were not entirely positive, however. Bjørn Stensaker's interviews at Swedish institutions indicated that authors of the unit-level self-evaluations often did not understand their purpose or how best to go about preparing them.

> Only a few interviewees apparently thought they and their departments profited from the institution's self-evaluation and the accompanying report. The majority of department heads are in other words not satisfied with these two elements. The reason they give for their dissatisfaction is that they felt the self-evaluations often became a passive response to externally defined needs, and that the reports were "weak" in that they did not specify utilisation of the self-evaluation for developmental objectives in their own department, and that the department heads were little involved in writing the self-evaluation report. [40]

There is nothing inconsistent about getting favorable reactions at the institutional level and skepticism at the departmental level. Such discrepancies would be expected given that the concept of quality work has yet to become established in all parts of the institution.

The instructions for preparing the self-evaluations generated a considerable amount of criticism. In Denmark, critics said that the performance indicator statistics were irrelevant, misleading, and hard to collect. Furthermore, the qualitative questions were too detailed, duplicative, in the wrong order, or insufficiently tailored to the subject being evaluated. These criticisms arose in almost every kind of institution and *uddannelse*. Respondents indicated that they would have liked more freedom to ask questions of their own, and they also expressed concern that much required information did not seem to be reflected in the steering committee's reports. Despite the sometimes sharp criticisms, however, the general feeling seems to be that the self-evaluation process did produce useful results.

Sweden's guidelines were criticized as providing too little direction, even though most respondents saw the permissiveness as positive. Some felt that while the lack of a clear outline or checklist of questions made the work heavier and more demanding, the process still turned out to be rewarding.

> *The instructions were not very clear, but this made us create a method of our own—which probably engaged us more.*
>
> *The discussions about what the self-evaluation should contain were rewarding and resulted in our completing the instructions with another method. Still, we missed reflecting on some important issues—examinations for example—which should not happen. Perhaps the National Agency should put up some obligatory criteria.*[41]

The Agency's own evaluation indicates that "[a]bout half the reports do not contain much reflection and analysis." Östling indicates that a chapter that "thoroughly discusses the issue of self-evaluation of quality work" might be a "relevant improvement" to the *Guidelines*. "Yet," she adds, "a detailed outline is not what the institutions would like and probably would not improve the reports."[42] I do not agree with her final point.

While the Agency's desire to fully empower the institutions is understandable, problems were inevitable given the rather primitive state of knowledge about quality work and its evaluation. Most academics have a reasonably good idea about program evaluation, but the self-evaluators of quality work had to invent their methodology as they went along. Östling's respondents indicated that the seminars on self-evaluation, formal talks with Agency officers, and speeches by the Chancellor helped a great deal. However, it appears that the concepts and terminology of quality work—including the essential distinction between quality work and quality itself—were not as firmly established as the self-evaluators might have wished. Furthermore, as one institutional respondent indicated, "The institutions were not sure whether self-evaluation was an instrument for improvement or the discharge of an obligation imposed by audit."

It appears that the unit-level self-evaluation reports were more valuable to the audit team than their institutional counterparts. These reports were not always included in the official self-study report, but they were made available on request. Some auditors went so far as to say that if they had to choose between the institutional and unit-level reports, they would choose the latter.

> *It was valuable to have the unit self-evaluations even though they were uneven in quality. When we had requested the unit self-evaluations and read them, we eventually got a fair picture of the current quality work.*
>
> *The institutional report was all too general whereas the unit self-evaluation reports gave a better basis for the audit.* [43]

Experience with Hong Kong's quality process audits produced exactly the same conclusion. The institution-level reports and presentations usually came across as rather pro forma. However, when institutional policies and rhetoric were juxtaposed with conversations and document review at the departmental level, the true state of affairs became readily apparent. The convergence between these two similar but entirely independent evaluation processes represents an important emergent conclusion about the methodology of audit.

The written record and interviews suggest a significant degree of variation in the process by which the self-evaluation reports were prepared. At the positive end of the spectrum, "the seven faculty members most responsible for teaching the subject divided the work among themselves and then held a two-day seminar with all teaching staff where the draft was presented and discussed. The same group also asked its external trade and industry council to comment on the draft." At the other extreme, a single person wrote the draft and then reviewed it with colleagues—"who were not so interested and did not have time" [44]—and some students I interviewed felt they had been excluded from the process. Doubtless the "lone wolf" approach was a rarity in both countries, but it appears that participation often was not as broad as expected. Both agencies might consider strengthening the guidelines for collegial self-evaluation, but in ways that are not disempowering.

The final question about the self-evaluation reports concerns their honesty. The Danish interviews yielded the following spectrum of responses:

> *Everybody is committed to the report. It makes it easier, now, to carry changes through because everybody has agreed on the weaknesses that we wrote about in the report.*

Or,

> *We had some different opinions about what really should be written in the report—everything we discovered in the process is not in the self-evaluation. The deficiencies often depend on individuals and those you don't dare to criticize anyway. That is why nothing happens, at least not as a consequence of that report.*
>
> *No, [the self-evaluation report] cannot be used. It is far too superficial, nothing is written about the real problems. That was a conscious strategy from the person responsible for the self-evaluation. We have other, internal documents, and we work with these instead.* [45]

Lack of forthrightness appears to be a pervasive problem in self-evaluation when assessments include accountability goals. Institutions and individuals may not see the benefit of disclosing weaknesses to third parties unnecessarily. Sometimes the formative benefits carry the day and a forthright report ensues. However, the process can become a game of limiting disclosure—say enough to avoid being caught by the site visit team but don't go into areas that are not likely to be probed effectively. The "private" conclusions can be used to effect change internally, but the improvement process is undermined by failure to confront all the issues during the self-evaluation and in the "public" document produced for the site visitors. Emphasizing the improvement agenda will mitigate these problems, but it seems they do not go away entirely.

Forthrightness did not seem to be as much of an issue in Sweden, although overtones could be heard. One auditor reported a discrepancy between what was written in the self-evaluation report and what was experienced during the site visit: "The self-evaluation report should have been more honest and analytic." [46] On balance, though, a careful reading of the documents and my own interviews suggest that the problems stemmed more from a lack of understanding about quality work and how to do a self-evaluation than from disingenuousness. For example, "The self-evaluation report gave a good basic description of the situation at the institution but was not applicable as a tool for the audit." [47] Stensaker supports this view. He reports that fear of misuse now is not much of a factor in either Sweden or Denmark, and that "the problem of honesty is more technical/methodological than strategic/cheating." [48] This reflects the situation I encountered in the Hong Kong quality process audits. I believe that more energy should be focused on these issues. For example, increasing the number of people involved in writing the self-evaluation reports and getting better guidelines should broaden the reports' perspective and sharpen their analyses.

There are several possible reasons why the Swedish self-evaluations appear less vulnerable to disingenuousness than their Danish counterparts. Because the self-evaluation guidelines were more general, respondents were less likely to get into areas where honesty would likely become an issue. In other words, audit's subject matter may make honesty less of an issue because one is inquiring about quality assurance and improvement *systems* rather than individuals' teaching performance. The Chancellor's strong emphasis on audit as a partnership between the institutions and the Agency also must have helped, as did the oft-repeated appeals to work together in developing the criteria for identifying and improving quality work.

Will audit's advantage with respect to honesty persist to the time when criteria are better established and accountability has inevitably become a larger part of the picture? I believe so. In addition to the advantages cited above, my experience with the Hong Kong quality process audits suggests that it is very difficult to disguise the quality of quality work. Respondents can't describe processes or performance measures that don't exist, for example, and site visitors have little difficulty identifying puffed-up descriptions.

User Surveys

Response to the Danish user surveys appears to have been mixed. On the one hand, they served a useful purpose by driving home the importance of stakeholders and providing new insights about their thinking.

> *The survey corresponds with the apprehension we have, so we have use for it. We are on our way to changing the deficiencies that the former students bring up. We were partly doing that before the evaluation. For example, the computer availability, about which students are complaining, is something we are improving now.*
>
> *You should of course not do all of what the employers say, but they are a very good start when you discuss how the uddannelse can and should change.* [49]

But the inevitable problems of survey design, sampling, response rate, and interpretation conspired to undermine the program's effectiveness.

> *Few of the employers who were interviewed could be <u>our</u> students' employers. Discussions with our own external trade and industry council give much more.*
>
> *The survey of former students was not very usable since so much has changed at the uddannelse after they graduated. It is to bad if it seems like we have defective computer education for example, when it is not so.*

> *It was obvious that most of the students that answered came from <u>one</u> part of the uddannelse, which slanted the answers.*
>
> *[The user survey] was for the steering committee and what was relevant in it is brought up in the report from the steering committee. That is why we have not used it or discussed it.*[50]

Institutions also cited the timeliness of completion as problematic. Sometimes results were not available to the steering committee before the site visit, and on at least one occasion they were delayed until after the committee had drafted its report. Some institutions wanted to see the results before preparing their self-evaluations—however, that might not be consistent with the evaluation's design philosophy.

The user surveys are innovative and consistent with the Centre's desire to focus institutions on the importance of stakeholder feedback. Experience confirms their value, but it also underscores the problems that arise when an external group conducts the surveys. It is difficult for such groups to get the survey details consistently right, and the departments may take any shortfalls as reasons to reject the results. It would be better, in my opinion, for departments to do their own surveys on a regular basis. Making their own decisions and interpretations would increase validity and buy-in. Perhaps the Centre could dispense with the cost and trouble associated with doing its own user surveys and ask the institutions, either separately or as a group, to do the job as part of self-evaluation—or better yet, as part of their ongoing quality work.[51] Methodological guidelines could be provided or Centre staff could participate in the planning process, but devolving the responsibility would eliminate disepowerment and enhance institutional learning.

Site Visits

The main elements of the visit programs appear to have gone smoothly. They were generally received positively in both countries, although some dissatisfaction can be discerned. Swedish institution-level respondents generally praised the panelists' knowledge and dedication, and the Agency's evaluation was positive as well. The visit structure evolved over time, which was consistent with the formative nature of the exercise. For example, one respondent indicated that the panel split itself into subgroups in some of the later-stage audits, and that this was very useful. The same respondent went on to say that at the large universities the teams rarely penetrated below the Faculty (i.e., school) level, but they should have gotten to the departmental level. One university criticized the process by saying "too much was left to the audit teams, which introduced uncontrolled variation—this was a very weak element." At another university, student representatives complained that the panel didn't take their inputs seriously enough.

Some department heads at Swedish institutions "expressed disappointment" about the visits of the audit team and their reports.[52] They cited the tight time schedule, the large number of people present, and the consequent formality of the meetings. Too much time was spent verifying facts, and on organization structure and administrative issues. Not enough time was spent on informal discussion of the department's quality work. This seems reminiscent of the UK (though not the Hong Kong) quality process audits, and it may result from a certain vagueness in the definition of quality work. As discussed later, the better the understanding of quality work the less the tendency to fall back on formalities. On balance, it appears that the site visit teams did learn a great deal at the departmental level and that this proved valuable in evaluating the institution's overall quality program. The program could be improved, however, by lengthening the site visits at large and complex institutions to permit a greater focus on the Faculty and departmental levels.[53]

Danish Centre respondents expressed considerable enthusiasm about the site visits. One went so far as to say that a strong steering committee could overcome weaknesses in the self-evaluations and user surveys. This positive view was echoed in Östling institutional interviews.

> *It was important to talk to the steering committee. It had a securing effect and was not unpleasant at all—the fears we had before it started disappeared after awhile.*
>
> *The visit was very positive. The steering committee was analytical. They asked reasonable questions without getting caught in details.*
>
> *It was interesting having them here and there was actually time for debate.*[54]

Two rather predictable kinds of criticisms about the visits did emerge, however. Some panelists were cited as being ill-prepared. (Busy people may not master all the material in the self-evaluation despite training and the best efforts of staff.) Some industry representatives "did not understand universities" and some academics from elite universities "did not understand universities with less traditional missions." Overall, though, it appears that the careful attention to the visit agenda and the selection and training of the site visit teams produced the desired salutary effects.

Hong Kong's experience reinforces two aspects of the Scandinavian reactions to the site visits. First, the site visitors seem to have less difficulty evaluating quality work than they do in evaluating quality itself. Second, departmental visits can provide better visibility than institutional- or school-level (i.e., faculty-level) meetings. This should come as no surprise. As Appendix A indicates, the main action in quality work takes place "hands-on" with teaching and learning at the departmental level. Institutional

and school meetings can evaluate quality work at those levels (which supports departmental quality work), but there is no substitute for sampling departments. Finally, the Hong Kong experience suggests that more time should be spent in the large and complex institutions than in smaller, simpler ones where there is less ground to cover.

The Reports

Reaction to the reports has generally been favorable in both countries, though there is an undercurrent of concern about their generality. Both agencies believe that while a certain amount of criticism is inevitable, the reaction has been positive overall. My interviews at the institutions indicate that at least the people most closely associated with the evaluation process feel the reports are on the whole fair and useful. More specificity would have been helpful, however, and the nature of follow-up is not clear. A Danish respondent indicated that the Swedish reports are much "softer" than their Danish counterparts, perhaps because of differences in team composition. (The Swedish teams consist mainly of nationals and include many university officers, whereas the Danish teams consist mainly of outside experts.) One Swedish respondent suggested that I ask the National Agency what it has learned from the stack of reports and what it plans to do with them.

Östling's interviews at Danish institutions document a number of grass roots reactions:

> *First we had a whole-day seminar where we laid the foundation for the subsequent development, where all employees took part. It was very positive. Those who still were a little bit negative to the evaluation changed their view during that seminar. Then we set up a coordinating group who sees to it that changes are carried out. To set that up was one of the best things we have done.*
>
> *I have used the report when I have been fighting with other departments about the evaluation—I then point to the report and say, "it says here that so and so is what we should do."*
>
> *No, you can't isolate what has happened—what's a direct case of the evaluation or what's not. Surely the report to some extent directs the work of change. But it would have been more visible as an active tool if we had gotten more criticism.*
>
> *We would have carried through some changes anyway, but maybe not so fast and not all of them. Certainly the evaluation had an effect. The greatest effect is the positive view we now have got to evaluation.*

> *We are using the report as an inspiration for further quality development but we are not uncritical to it—the recommendations are not answer lists. Some administrative recommendations that the steering committee gives they don't understand the meaning of and for that reason they are impossible to follow.* [55]

These accounts suggest that the while the reports could have been more specific they did include actionable elements.

Bjørn Stensaker reports similarly mixed departmental reactions to the Swedish audit reports.

> Our group did not feel that the reports were very important for them, and that the department and/or institution's own strategies and thinking about quality improvement were of more use. The other group [of about equal size] viewed the external reports as a small but important part of a larger developmental process. [56]

He goes on to say that "the critical attitudes towards the auditors' [departmental] visits and their reports do not appear to have resulted in a rejection of quality improvement work and the auditing process." [57]

The Swedish departmental visits—and presumably their coverage in the reports—came about as something of an afterthought and were not as well organized as other parts of the audit program. This may account in part for the negative reactions. Moreover, the *prospect* of an evaluation and the discussions with the site visitors may well have spurred development to the point where the audit report contained largely "old news." But despite these caveats, the criticisms call out the need to make all parts of the audit reports as operationally meaningful as possible.

Wahlén reports that on occasion the response to the Swedish reports was genuinely enthusiastic even when they included severe criticism.

> That particular institution argued that the report gave management the strength to pursue policies, which would otherwise have been accepted only reluctantly. One rector expressed the opinion that the visiting team should be transformed into a permanent advisory group. After all, there was no other group which knew the institution and its strengths and weaknesses so well. [58]

He goes on to say that,

> The audits have been commented on favorably in internal staff magazines, with indications that these are the areas, which are now at the focus of the rector's attention. [59]

I have encountered similar expressions in Hong Kong and in the UK. Quotations of the first kind are typically heard in private conversations with senior university officials or quality enthusiasts who have finally succeeded in gaining the ear of such officials.

Section Five: Impacts

The reactions to evaluation suggest that the processes did indeed affect behavior. My analysis begins by describing the post-evaluation level of quality work. While one would like a "before" measure as well, it's clear that quality work was in a fairly primitive state prior to the onset of evaluation. Next come my assessment of the impact of the two evaluation programs on quality work, a discussion of how evaluation affects academic autonomy, and some general comments about the diffusion of quality work across institutions and departments. The report ends with suggestions for improving the definitions of quality work and the methodology of quality audit.

Current Quality Work

The best quality work in Danish and Swedish institutions appears to operate at or near the state of the art. It covers all five domains described in Appendix A: design of curricula, teaching and learning processes, and outcome measures; the assurance and continuous improvement of implementation quality; and communication of exemplary practice across schools and departments. It also covers all six "essential principles": defining quality in terms of outcomes, focusing on teaching and learning processes, striving for coherence, working collaboratively, basing decisions on facts, and striving to emulate exemplary practice through benchmarking. Perhaps most important, the principles are applied self-consciously and systematically. Both the institutions and the evaluation agencies cite organizational learning as a major objective, and many quality work routines have been organized with this in mind.

Appendix B presents some exemplary practices gleaned from my interviews and reading. By "exemplary" I mean practices that illustrate the evolving concept of quality work and which may be worthy of emulation. They are not necessarily "best practices," and I don't wish to suggest they cannot be found in other institutions and departments—including departments at the institutions I visited.

Exemplary practice in Sweden and Denmark appears comparable to the kind Wilger and I have observed in Hong Kong, the UK, Australia, and a few U.S. institutions. (Wilger will visit continental Europe and the UK this fall.) While our search has not been exhaustive, we think we have a reasonable grasp of the state of the art. What I saw in Sweden and Denmark sits squarely in the domain of good practice and in some cases

rises to the level of best practice. For example, the "System for Quality Assurance and Quality Development" put in place by Professor Bengt-Ove Boström and his colleagues in the Department of Political Science at Göteborg University and reproduced as Appendix C, is the best such program we've seen. The two countries' performance is all the more remarkable because the system was highly traditional, with little if any quality work, just a decade ago.

Despite the impressive progress, quality work in Sweden and Denmark remains undeveloped relative to its potential and vulnerable to setbacks—as indeed is the case in all countries. Exemplary practices occur with considerable regularity, but they are far from common. My sample did not encompass a cross section of practice (I chose to visit exemplary institutions and talk to representatives of exemplary departments), but the evidence suggests large disparities in performance among organizational units. Quality work is at an early stage of diffusion worldwide, and it has yet to be universally accepted even in Sweden and Denmark. Still, these two countries can justly claim to have reached the frontier of development.

Effect of Evaluation on Behavior

To what extent can evaluation be credited with the advance of quality work in Sweden and Denmark? In my opinion the answer is, "To a very significant extent." The evaluation programs were triggered by public policy reforms that decentralized responsibility for quality while maintaining accountability. Regulation was replaced by a philosophy of "soft managerialism," which extended from government to institutions' central administrations to faculties to departments. As one of my respondents put it, "Trust, but check." And by checking, one also steers. I believe that evaluation stimulated quality work across both countries' higher education systems to an extent that would have been inconceivable otherwise.

Participants in the evaluation process describe its effect as starting a discourse about quality and quality work. I agree. Academics can produce remarkable results when they turn their minds to an issue. The biggest inhibitor of quality work is lack of attention—too little time on task. Next comes the tendency of faculty to work as individuals on most education-related matters, what Wilger and I have termed "hollowed collegiality." Starting a quality discourse represents the necessary first step toward mitigating these difficulties.

The importance of discourse can be illustrated by noting that, in 1993, the deans of many Swedish faculties fought *any* institutional quality initiatives. Many professors felt they knew all there was to know about quality, and they resented suggestions that they might do things differently. Developing a vocabulary with which to discuss quality and

quality work represented a necessary first step, a prerequisite for building a case for improvement. Then it was necessary to build a common consciousness, which in turn could produce a stimulus for change. Providing venues for discussion, often in the form of faculty and institution-level committees or "quality councils," turned out to be very useful—provided that the councils developed a coherent agenda and received high-level backing. The case of Uppsala University, described in Appendix B, illustrates how the process worked. The aforementioned quality program of the Department of Political Science at Göteborg University (Appendix C) illustrates a best-case outcome.

Unfortunately, discourse doesn't always produce the desired result. Many faculties and departments have yet to internalize the principles of quality work, and some continue to actively resist the idea. (One respondent indicated that only 20 percent of departments were actively engaged in quality work. While that may well be high by international standards, it does indicate substantial room for improvement.) Future quality audits might put greater emphasis on internal and external benchmarking. While recognizing that every entity should be able to choose its own quality work program, one can ask why some units put more effort into the process than others. Benchmarking disseminates information about exemplary practice and leads to "why not?" kinds of questions. Surely each academic unit should ask such questions, even if an informed analysis leads them to their own unique conclusions.

Bjørn Stensaker's review of the Swedish evaluation reports describes another kind of discourse—one between the institutions and the broader society.

> [The audit reports] represent a focused dialog—a discourse—about quality between representatives of Swedish society and the higher education institutions in the country: a discourse that when analyzed can indicate how society recommends and preconceptualises the changes that should take place at the institutional level.[60]

He goes on to say that while the discourse may elicit strong disagreements about what changes should take place, the societal views may be difficult for institutions to reject. Universities exist within the society and are largely dependent on it for resources. Both neo-institutional theory and resource dependence theory predict that institutions will adapt as needed to protect their resource base.

The majority of Swedish audit report items concern "institutional objectives and strategies" and the "management and organization of quality improvement work."[61] Because they focus heavily on quality work, the reports—and the evaluation process that led up to them—represent a large burst of energy aimed at eliciting change. Without such an energizing effect, the growth of quality work would have been much slower.

The reports also placed the responsibility for instilling change more firmly in the hands of institutional leadership than had ever been done before. Some professors expressed surprise that the touted devolution of responsibility from government to academe did not free the faculty from all accountability. The audits made it clear that with freedom comes responsibility, and that society expected the institutions to further its quality work agenda.

The strength of the Scandinavian evaluation philosophy lies in its strategy of creating discourse rather than laying down regulations. We have seen how, in the right circumstances, this can produce real change. Implicit in the approach—and explicit in the case of Sweden's retreat from formal linkages to funding—is that the discourse should indeed be two-way. Society (as embodied in the evaluation agency and its expert panel), institutional leaders, institutional quality advocates, and rank and file faculty all press their case. If managed effectively, the process can achieve constructive resolution and an advance in quality work performance. Regulations and evaluations that are heavily slanted toward accountability cut off discourse before it starts and turn the exercise into an adversarial game.

Effect on Academic Autonomy

Questions about academic autonomy arose early in the evaluation process, especially in Sweden. Evaluation came forward as part of a devolution program that put more power in the hands of institutions, and by implication faculties and departments. Critics of evaluation worried that it would be unduly constraining, either in terms of specific processes or in terms of institutional governance generally.

Bjørn Stensaker poses the question of *institutional* autonomy this way.

> Higher education institutions may not only change their organizational structures in similar directions (isomorphism), but perhaps also their main activities—teaching and research. A possible long-term effect of such a change could be that institutions develop to be more homogeneous and standardized, and that institutional uniqueness and diversity may be lost on the way.[62]

Concerns about uniqueness and diversity are heard frequently in the United States, so these questions are of special importance for our research.

Analysis of the audit reports provides little support for isomorphism. The Swedish recommendations have been described as "often quite abstract and general, referring to a broad objective or ideal conditions without…specifying how the institutions should act to implement [them]." Some recommendations are so broad as to defy specific

follow-up. "What is actually meant by, for example, 'developing the leadership role'? Should this be understood as a need to improve the conditions for management/leadership, or as a need for leaders to steer more?" [63] Such broad recommendations hardly represent managerialism, and they are not likely to stifle uniqueness and diversity. (As will be discussed later, the abstract character of the reports presents its own problems.) The Danish reports are similarly broad, and although Denmark appears to be more centralized because of its *uddannelse* structure, it is hard to imagine them as a significant force for isomorphism.

Professor Berit Askling, head of the quality unit at Göteborg University, describes the faculty's concern that, by holding institutional leaders accountable for quality work, evaluation would undermine internal academic autonomy.

> Among academic staff and heads of departments, the internal attempts to strengthen institutional leadership appear as a paradoxical consequence of a decentralizing reform and are usually regretted. For these categories, the reform "of freedom" has led both to some loss of individual freedom and a decreasing space for collegial decision making, and there is also a tendency for disassociation between institutional leadership and faculty. The institutional leaders are sometimes seen as bureaucrats, more closely allied with state authorities than with their own staff, something that might motivate one to categorize them as a new "intermediary body." [64]

Askling disputes John Brennan's view that, "At the macro level, quality is about power and control." [65] At the same time, she does not give up the idea that quality work can leverage institutional policy.

> We can say that, so far, neither the power aspects nor the control aspects have been predominant in the management of the quality issue. This is true at both the national and the institutional level. However, at the institutional level, the quality issue has, at least at some universities, given the rectors incentives to establish interfaculty bodies and, in so doing, strengthen the institutional leadership. [66]

The so-called "quality councils" demonstrate that quality work need not be an exercise in power and control (The term comes from business, where in the 1980s firms like Hewlett Packard used it to distinguish professional-level quality work from the then-fashionable shop-floor "quality circles.") Many institutions created such councils in order support quality work, but they were careful to eschew evaluation. In Askling's words:

> The quality councils did not assume the role of an executive assessor of quality, as such, of the institution, or the students' outcomes, or the institutional management. They have tried to distinguish themselves mainly as supportive bodies and to respect the responsibilities of the faculties and departments for the more explicit defining and assessing of quality.
>
> At one of our universities, the first quality council was expected by the university board to undertake evaluative tasks, but refused to execute them and, consequently, asked to be dismissed.[67]

My own interviews abundantly corroborate Askling's conclusions.

While the concern about autonomy was heard frequently, the reality has proven much less threatening. Most academic leaders adopted the kind of "light touch" that characterized the external evaluations themselves. Most faculties and departments remained free to develop their own strategies for quality work. As several respondents put it, the only requirement was that the unit *have* a quality work program. While even this might have rubbed some traditional academics the wrong way, it hardly represents a bureaucratic straitjacket.

Has the process of evaluation stifled uniqueness and diversity? I think not. For one thing, decentralization is an established fact in most Swedish universities. While evaluation has shifted a certain amount of responsibility from faculties and departments to the central administration, most initiatives remain in local hands. The central programs have been carefully designed to stimulate, not replace or regulate, work at the local level. This lies at the heart of the philosophy of discourse, and in any case the faculties and departments could find ways to veto anything that might be imposed on them. The expressed purpose of evaluation was to increase professionalism in the area of quality work, not to induce a "managerialistic" culture. This includes professionalism at the local level as a central tenet. One expects, and indeed observes, that local initiatives will further uniqueness and diversity.

The Diffusion of Quality Work

Despite their salutary effects, it would be a mistake to say that the two Scandinavian evaluation exercises transformed higher education. Universities are too complex, their governance systems too convoluted, and their faculties too traditional for real transformation to take place in only a few years. The Swedish tradition, moreover, celebrates the autonomy of individual academic units within the universities, which makes transformation particularly difficult to achieve quickly.

What evaluation sought to do, and to a large extent did do, was to start universities and departments on the road to becoming learning organizations—not learning in the sense of traditional scholarship, but rather to become self-conscious about the processes of teaching and learning and how to improve them. Askling puts it this way:

> Quality work can be used as a tool for promoting a department towards a learning organization. However, my feeling is that there is still a long way to go before we will be able to use this tool efficiently. Not just the role of leadership but also the internal devolution of authority must be considered. In addition, the supportive and controlling aspects of quality work must also be taken into account and used by the institutions as tools for promoting a collective awareness and a better self-understanding.[68]

Evaluation represents an intervention, by external authorities, to accelerate the diffusion of quality work. Quality work, in turn, aims to promote self-understanding about teaching and learning processes and instill an intrinsic and collective desire to continuously improve them.

Stensaker's review of the Swedish evaluation system concludes with this caveat about the timeliness of audit given the then-current degree status of quality work.

> The auditing processes seem to be more addressed to the structure of the institutions "such as they should look in the future," than the way they actually are today. Of course, audits are tools for change, which can contribute to fulfilling the intentions behind the 1993 reform on stronger institutional management. However, it may be asked whether today's auditing processes are not too far in front of realities.[69]

Certainly audit's controlling aspects were too far in front of realities—departments couldn't be controlled and there was little quality work to audit. However, viewing audit as an intervention aimed at furthering the diffusion of quality work leads to a different conclusion. The evaluation systems were designed to infuse a new kind of energy into the higher education system, to make the universities and departments learning organizations in the sense described above, and in this they were strikingly successful.

Consider the matter from the government's point of view. Having identified a gap (in this case concerning quality work) between the objectives and performance of the universities and those of the larger society, leaders must choose between two alternatives. They may sit back and wait for more propitious times, when the universities will have embraced the new objectives of their own accord, or they may become proactive in trying to elicit change. Many academics would prefer that the system be left alone, but

stewardship and political realities argue for intervention—especially when public funding is an issue. The key is to intervene without intrusive regulation or micromanagement. The "light touch" adopted in the Swedish and Danish evaluation exercises provided the needed interventions without violating institutional autonomy.

The diffusion of quality work follows the s-shaped curve associated with the adoption of innovations. The process is slow at first, then accelerates as innovators give way to early adopters and eventually to the majority of users. (Growth eventually slows down as the population of potential users is exhausted.) Figure 3 of Appendix A shows the s-shaped curve and describes the various stages of adoption. A potentially valuable innovation may fail to take off into self-sustained growth because the majority of potential users don't recognize its advantages—either because the case isn't made effectively or because other things distract them.

Quality work in Sweden and Denmark now has become well established as a valuable innovation, and some institutions even have reached the stage of early adoption. However, it does not appear to have reached the critical mass needed for self-sustained growth. Some organizational units have reached the level of international best practice, and a significant number of others have reached exemplary status. But many units have been largely untouched and others seem to participate reluctantly. It also appears that quality work does not command the full attention of many institutional heads. They support the idea but do not yet see it as a key results area. One senses that many within the academy would be just as happy if the quality work idea were to fade away so they could get on with business as usual. This is not an indictment—indeed, the situation in Scandinavia is better than in most other countries—but rather a statement of fact with implications for the future.

Continued injections of energy will be needed to move quality work up the adoption curve, to diffuse it to the point where its advantages are self-evident and its use by institutions and departments is ineluctable. A loss of momentum now would jeopardize all the progress made so far. Quality work has grown dramatically but it has not yet permeated the academic culture to the point where it can compete on equal footing with the traditional emphasis on teaching and research.

Evaluation should not only be continued, it should also be increased in rigor—with more specific definitions of quality work and what is expected of institutions and departments. The abstract character of the quality work, which helped mitigate the initial worries about institutional autonomy, will tend to inhibit further diffusion. Whether quality work will become firmly established depends on whether the evaluation processes can be refined and carried forward without loss of momentum.

The Danish quality system was reaffirmed in June 1990 when Evaluation Centre was integrated into a new Danish Institute of Evaluation. The Institute will undertake systematic and mandatory evaluation of teaching and learning at all levels of the Danish educational system—from primary school through post-graduate programs. This will extend the scope of evaluation and enable the seamless review of teaching in a discipline from the primary to the tertiary level.[70]

The Swedish National Agency's responsibilities now extend well beyond the audit program, and there appears to be some danger that the priority placed on quality work may be diluted. They plan to continue the audit program for a second round, however, and the Agency's capacity to innovate remains in place. One hopes that the needed impetus and support will be forthcoming.

Conclusions

The Danish and Swedish evaluations focused attention on quality work and launched its diffusion into institutional routines. By soft pedaling the accountability agenda, they opened the way for serious discourse, which if pursued is likely to produce major dividends in the years ahead. By carefully designing the processes and implementing them with a spirit of openness, the two quality agencies earned the trust of the institutions while simultaneously challenging them to improve their quality. Those familiar with evaluation know that such outcomes are by no means automatic.

Whether this has improved the delivered quality of teaching and learning cannot really be demonstrated. However, there are strong prima facie reasons for believing that quality will improve if it has not done so already. Defining quality in terms of outcomes, focusing on teaching and learning processes, striving for coherence, working collaboratively, basing decisions on facts, and striving to emulate exemplary practice through benchmarking certainly should produce improvements. Experience outside higher education indicates that failure to do such things produces a high risk of mediocrity. Wilger's and my research for NCPI demonstrates the applicability of this experience to higher education. Furthermore, we can find no "invisible hand" that produces the desired effects without well-organized and intensively pursued quality work.

The Scandinavian experience also demonstrates the efficacy of academic audit—that is, the systematic review of institutional and departmental quality work. Sweden adopted audit as its methodology of choice and pursued it diligently across the higher education system. Denmark chose subject-level assessment, but included the review of quality work as a central element of its process. Both countries advanced the state of the art in audit, as well as in quality work itself. They have shown that audits need not be overly

expensive or intrusive, and that they can successfully stress improvement as well as accountability. My own work with Hong Kong's quality process audits strongly supports these conclusions.

There are several reasons for believing that academic audit dominates subject-level assessment on a number of important dimensions. Audit focuses on quality work, which is proving to be easier to evaluate than quality itself. The case that effective quality work is a necessary condition for optimal quality is growing in strength, and observation indicates that such work is lacking in most academic organizations. Whether effective quality work represents a sufficient condition for optimal quality is debatable, but there seems to be little doubt that improving quality work will improve quality, other things being equal. The Swedish and Danish experience illustrates how academic audit methods can significantly improve quality work.

Respect for institutional uniqueness and diversity provides a second reason for including audit in any evaluation program. Assessment requires criteria that define a quality education for each subject area and institution. There are two choices: the evaluating body either must set its own criteria or it must rely on criteria established by the institutions themselves. Centrally determined criteria undermine uniqueness and diversity, whereas local criteria require that purposes and performance be specified clearly—which is one of the elements of academic audit. One also should ask how the criteria and indicators were developed, how they are interpreted and used, and whether they are being continually improved. Once again these are academic audit questions.

Suppose a program passes all the aforementioned tests: well-qualified faculty work collegially to make fact-based design decisions in light of well-elaborated criteria, they continually evaluate learning outcomes, and they refine the curriculum and teaching methods in light of the feedback received. What assessor would substitute his or her opinions for such careful and collegial judgments? Would this be appropriate? Assessors should ask whether the local process is good and whether decisions flow logically from the evidence, but they should not second-guess the considered judgments of those on the scene. As in Denmark, the effective assessor will be asking audit questions much of the time.

The "topology" of the higher education system may provide another impetus for audit. Most systems consist of relatively full-line institutions, so that the typical subject-level assessment will encompass most of the schools. If there are n institutions and m subjects, the number of assessment visits will approach $n \times m$, while the number of audits would only be n. The UK can attest to the difficulty and expense of mounting $n \times m$ visits in a large full-line system. In Denmark, where only a few institutions are full-line, the number of visits proved manageable—indeed, it may have been easier to run a

series of small self-contained assessments than a single audit exercise covering all twelve universities or the large number of sub-universities. Subject-level assessment allowed the Centre to examine the quality of students, faculty, infrastructure, and similar factors not included in academic audit. The approach also invited detailed scrutiny of performance indicator trends and benchmarks. On the other hand, it did not provide visibility into institutional quality work above the level of the *uddannelse*.

Audit suffers from one major shortcoming as it has been practiced so far: the lack of sufficiently detailed definitions about what is meant by "quality work." Respondents cited this problem in their discussion of self-evaluation. Absent clear definitions covering the totality of quality work, evaluation teams tend to fall back on its administrative and organizational aspects. The lack of clear quality work definitions also makes it harder for the evaluation reports to offer operationally meaningful action recommendations.

The UK tried to provide specificity by focusing on the administrative aspects of quality work: for example, are there written quality assurance procedures, are they followed and documented? Unfortunately, an administrative emphasis gives the whole exercise a rather bureaucratic cast. In any case, it misses the core elements of quality work—the faculty's direct professional efforts to improve and assure quality. No wonder many UK academics feel that quality audit is a rather sterile exercise.

The Swedish and Danish evaluation agencies understandably wished to avoid over-prescription but, in retrospect, a more detailed description of quality work would have been desirable had one been available. Once again, the Hong Kong experience mirrors that of Scandinavia. All three exercises required "learning on the job." Producing more detailed descriptions of quality work would have been difficult if not impossible in the 1994-6 time frame even if concerns about over-prescription could have been surmounted. Definitional generality probably helped win initial acceptance of the quality work idea, but further diffusion will require a greater shared understanding.

The more detailed descriptions should provide concrete guidance about the nature of quality work, how to go about doing it, and how to "know it when you see it." On the other hand, uniqueness and diversity require that the descriptions stop short of telling institutions and departments exactly what they should do. Such micromanagement would erode the benefits of what Burton Clark calls the "self-regulative university."

> Self-defining, self-regulating universities have much to offer. Not the least is their capacity in difficulty circumstances to recreate an academic environment.[71]

It would alienate what Frans van Vught calls "the academic heart" and its passionate search for excellence—and thus prevent true long-term quality improvement.[72] Inspired by my visit to Scandinavia and by Hong Kong, I am working with Andrea Wilger to develop quality work descriptions that can provide positive guidance without alienating the academic heart.

Appendix A, "Quality Work," presents our conclusions as they have developed so far. I believe the descriptions pass successfully between the Scylla of overgenerality and the Charybdis of micromanagement. That, of course, is for our readers to decide. We ask that the following criteria be kept in mind. Do the descriptions represent operationally meaningful propositions that offer reasonable prospects for improving educational quality? Can one think of an institutional type where behavior consistent with the propositions would not improve quality? (Quality should be defined as value added, not in terms of inputs or prestige.) Do the propositions constrain uniqueness and diversity, other than the undesired diversity of optimal versus lackluster quality? Could the propositions be cited in the instructions for academic audit, and would this be useful? The two of us would be grateful for feedback on these and any other matters covered in this report.

Appendix A: Educational Quality Work (NCPI)

William F. Massy and Andrea K. Wilger
National Center for Postsecondary Improvement
Stanford University, October 1999

"Quality work" has become a term of art in a number of countries, where it is used to distinguish between higher education quality assurance and improvement and the processes associated with the actual delivery of teaching and learning. This Note maps the quality work idea into the emergent conclusions of our research at the National Center for Postsecondary Education. That research also considers the linkages between educational and research quality, but we do not address such questions here.

Definition of Quality Work

Educational "quality work" refers to the activities of faculty, academic leaders, and oversight bodies that are aimed at improving and assuring quality. Quality work supports the provision of educational quality, but the two represent distinct kinds of activities. Quality provision depends on curricula and teaching and learning processes. Quality work focuses on the *improvement* of curricula and teaching and learning processes, and on *assuring* the best performance possible given available resources.

Quality work begins with the academic activities of faculty—for example, in program- or department-level teams charged with curricular improvement and finding better ways to teach, to stimulate learning, and to evaluate learning outcomes. It continues at the school and institutional levels, where committees or quality councils support, stimulate, and on occasion evaluate department-level quality work. External bodies may contribute to quality work by stimulating and evaluating institutional efforts. We shall call this "quality oversight" to distinguish it from the quality work that takes place within institutions, but quality oversight should be included in higher education's overall quality work agenda.

The evaluation of educational quality depends on subject matter, student characteristics, and institutional mission. Cross-institutional minimum standards can be established for purposes of accreditation, but quality levels above the minimum should be context specific. Quality work, on the other hand, can be evaluated using criteria that are broadly similar across institutions and subjects. Exemplary quality work can be identified and benchmarked wherever it occurs.

Excellence in quality work requires a high degree of collegiality and professionalism, and balanced priorities for teaching and research. Such excellence requires that professors work together rather than as individuals and that they devote substantial time to, for example, explicating educational goals, enhancing teaching and learning processes, and implementing performance measurement schemes. Our research shows that few universities' quality work programs are sufficiently well organized and systematic to maximize the benefits for educational quality. Likewise, few quality oversight programs focus on quality work to the extent needed to change institutional and faculty behavior.

Effective quality work cannot guarantee educational excellence, but a growing body of evidence indicates that it is a necessary condition. For example, well-qualified and dedicated faculty working according to good processes will produce better educational quality than counterparts who are inhibited by poor processes. We often remind ourselves that educational excellence cannot be achieved without sufficient numbers of qualified faculty supported by adequate resources. Effective quality work should be added to the list.

Domains of Quality Work

Educational quality work applies to five broad domains of activity.

(a) *design of curricula*: what will be taught, in what order, from what perspective;

(b) *design of teaching and learning processes*: what methods will be used, how students will learn, and who will be responsible for each educational element;

(c) *design of outcome measures*: how student learning will be assessed and its consequences determined;

(d) *implementation quality assurance and improvement*: how well faculty and other responsible parties implement the designs and work to improve their performance; and

(e) *communication of exemplary practice*: whether there are effective programs for diffusing exemplary practice across academic units; whether the institution or system is a "learning organization" with respect to educational quality work.

Traditional academic quality processes address item (a) by requiring committee approval for courses and to some extent for course content, but the approvals are mostly discipline- rather then student-oriented and rarely involve deep analysis. Individual professors sometimes address item (b), but because they usually work in isolation, wide-ranging innovation and organizational learning lag. Item (c) rarely extends beyond traditional grading processes. Item (d) is mostly limited to student course evaluation surveys, which tend to be heavily discounted except in extreme cases. Item (e) rarely gets the attention needed to make quality work a high priority and the institution a learning organization.

Essentials of Quality Work

The quality literature identifies six principles for effective quality work. It treats service suppliers as well as manufacturers, and nonprofit as well as profit-making organizations. The six principles represent quality work's essential professional core.

Our research indicates that the principles apply to all kinds of colleges and universities. We invite our readers to ask whether one could seriously argue against any of the six as a matter of principle.

The six principles are as follows:

1. *Define educational quality in terms of outcomes.* Shift from a teaching to a learning emphasis. Learning outcomes should be judged in relation to student needs, not in terms of tradition or strictly discipline-based criteria. Quality definitions should emphasize value added by the educational process to avoid confounding with incoming student diversity.

2. *Focus on the process of teaching and learning.* Education represents transformation, and transformation requires process. Process design is important, and so is the effectiveness of implementation. Changed circumstances often require process adaptation, which may trigger significant redesign.

3. *Strive for coherence in curricula and educational process.* View education as a continuous process, and make sure the parts interconnect seamlessly. Avoid the tendency to treat one's own part of the process as if it were a self-contained "silo."

4. *Work collaboratively to achieve mutual involvement and support.* Collaboration applies a broad range of skills and experience to difficult problems, and organized teamwork provides impetus for collaboration. Teamwork also implies collective accountability and encourages peers to hold each other accountable for individual performance.

5. *Base decisions on facts wherever possible.* Invest time and effort in the collection and analysis of data—for example, on student needs and learning outcomes—and organize to use it well. Document the assumptions and logic behind decisions when evidence is not available.

6. *Strive to emulate exemplary practice.* Identify and benchmark such practices both inside and outside the institution, then adapt these practices to local circumstances.

The six principles can help quality workers organize their thoughts and identify opportunities for improvement. Because it is not difficult to audit conformance, the principles also can guide evaluators of quality work. Yet they are sufficiently general to avoid infringing on institutional and departmental autonomy. The principles should be applied self-consciously and systematically to maximize organizational learning and embed good processes in organizational routines. Participants should strive for continuous improvement no matter how good their current performance.

Examples of Quality Work

We are compiling an inventory of examples to illustrate the concepts and the relevance of quality work. A brief sampling follows. Each activity is in regular use by exemplary higher education entities with which we are familiar.

- *Stakeholder surveys:* mail, telephone, or focus group interviews with employers or alumni. Stakeholders can provide information relevant to all five quality work domains. The surveys help define educational outcomes (principle 1) and further fact-based decision making (principle 5). They can fruitfully be incorporated into any level of quality work, but they probably are most useful when conducted regularly at the department or program level.

- *Value-added performance indicators:* assessment measures for learning outcomes (domain c). Such indicators may take the form of criterion-referenced tests, skill assessments, or unobtrusive observations of learner behavior. They are

best developed and used at the departmental or program level, where they help define educational quality in terms of outcomes (principle 1), highlight educational coherence (principle 3), and further fact-based decision making (principle 5). School and institutional quality work, and educational quality oversight, should stimulate and support local performance indicator development rather than supplant local with central measures.

- *Benchmarking programs:* systematic identification and evaluation of educational provision and quality work in other departments and institutions, followed by efforts to adapt their best features to local conditions. Benchmarking focuses attention on exemplary practice (principle 6) and applies to all five quality work domains.

- *Faculty teams with cross-disciplinary charters:* working parties that are empowered to design and implement change with a minimum of outside interference, subject to ex-post accountability. (Teams should be distinguished from committees, which typically determine policy or regulate the work of others.) Collaborative work teams (principle 4) can be used effectively in all five quality work domains. Cross-disciplinary teamwork also improves coherence (principle 3).

Prerequisites for Quality Work

Quality work requires strong leadership and a supportive rewards and incentive system. Presidents, provosts, deans, and other academic leaders should put quality work at the core of their visions and strategic plans, and support it vigorously with both words and actions. The faculty incentive and reward system should celebrate and reinforce educational quality work, and never subordinate it to other objectives. (Research incentives will coexist in some institutions, but they should not eclipse the incentives for educational quality work.)

The prerequisites represent important litmus tests for quality work—tests which internal and external parties can evaluate without much difficulty. Whatever the level of quality work within the institution, it can be improved by strengthening leadership support and providing better rewards and incentives.

Facilitators of Quality Work

Our research has identified many kinds of actions that can facilitate quality work performance. A partial list follows.

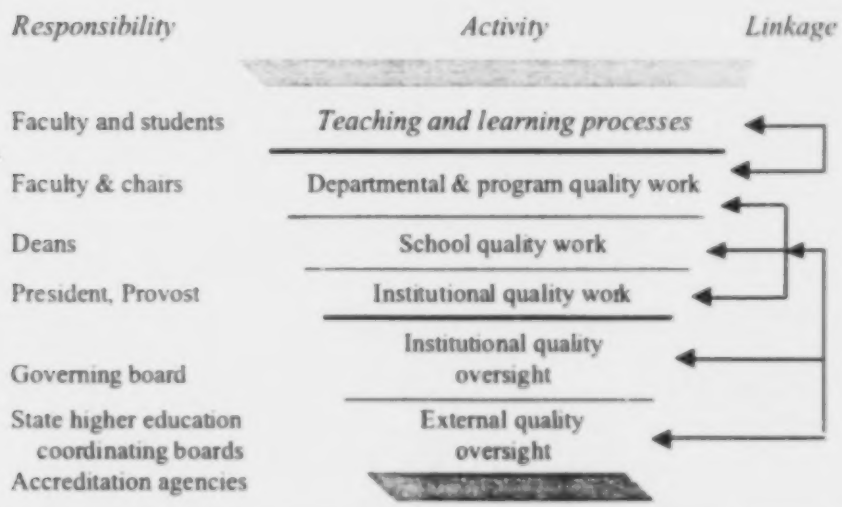
- clearly assigned responsibility for quality work at every level
- organized venues, sometimes called quality councils, to focus attention and discourse on quality work
- human resource development at all levels, including the instruction of doctoral students, with respect to quality work principles and practices
- internal advocacy and consulting centers for educational quality work
- short-term project funds for high-leverage quality work improvement
- positive linkages between quality work and ongoing budget-making criteria
- periodic audits of quality work performance

Audits may be performed by external agencies as part of quality oversight, by the institutions themselves as part of their quality work, or both. Audits performed by institutions may benefit from external inputs—for example, the inclusion of visitors from other institutions or industry.

The Quality System

Quality work takes place at a number of different levels inside and outside the institution. The "pyramid" shown in Figure 1 depicts the elements and interactions that make up the higher education quality system.

Figure 1. The Higher Education Quality System



The pyramid contains three major regions. *Quality work* within the institution sits in the middle, *teaching and learning processes* at the top, and *quality oversight*, at the bottom. By inverting the pyramid and flattening its bottom, the graphic calls out that quality work provides the foundation for teaching and learning quality. Putting the teaching and learning processes at the top and in larger type emphasizes that they represent the quality system's reason for being.

Quality work takes place mainly at the department and program level, where it interacts strongly with the processes that deliver teaching and learning. This interaction can be viewed as "hands-on" because quality provision and local quality work involve substantially the same people. Such close interaction is essential because teaching and learning are so complex and vary so much among departments, programs, and institutions. Effective quality work requires more than professionalism—it requires a change of culture. To borrow the phrase coined by our colleague Frans van Vught, quality work should "engage the academic heart" in a passionate quest for improved educational quality.

Local quality work combines feedback with problem solving. Gaps between performance and expectations lead to problem solving, and good performance should boost expectations. Quality work produces a virtuous circle of rising aspirations and performance.

School- and institution-level quality work support and stimulate departmental-and program-level work. Providing leadership, resources, incentives, information, training, and interdepartmental venues for discourse on quality work offer examples of school and institution-wide activities. Such activities also should include periodic evaluation of grassroots quality work to ensure accountability and provide impetus for improvement.

Quality oversight is designed to energize quality work and assure its effectiveness. Oversight begins with the institution's governing board. By establishing the quality work agenda and monitoring progress, boards can stimulate improvement without micromanagement. Audits by state higher education coordinating boards and accreditation agencies can do the same. By focusing on quality work, institutions and oversight bodies can discharge their accountability obligations without resorting to disruptive assessment practices or intrusive regulation.

Quality Evaluation

Worldwide interest in "academic quality evaluation" has been growing for more than a decade. The term includes "assessment," which means the evaluation of teaching quality, learning quality, or both. It also includes "audit," which means the evaluation of quality work. Research evaluations are important as well, but they fall outside our current scope.

Outcome assessments are critical elements of quality work because they provide the feedback needed for improvement. While professors often complain that assessment methodologies are too primitive, context-specific, or costly to be useful, such concerns apply more to externally mandated assessment vehicles than to those operated hands-on by departmental and program faculty. We believe that outcome assessment should be encouraged in local quality work but that centralization of assessment methods and criteria should be discouraged.

Audits of quality work, on the other hand, work well at any level of the quality system. Audits are not difficult and need not be costly or intrusive. Our research indicates that they can combine improvement and accountability objectives to an extent not possible with assessment. Refinement of the quality work concept will make audit even more effective. Looking at quality work systemically improves both evaluation and performance in a self-reinforcing way.

Martin Trow offers the fourfold typology for quality evaluation shown on the left side of Figure 2. He argues that supportive reviews initiated by institutions (Type I) will produce the greatest improvements in teaching and learning. Conversely, externally driven evaluative reviews (Type IV) produce little academic value and may lead to evasive strategies.

Figure 2. Quality Evaluation

Fourfold Typology

Origin of the review:	Function of the review	
	Supportive	Evaluative
Internal	I	II
External	III	IV

Martin Trow, *Academic Reviews and the Culture of Excellence* (Stockholm: Kanslersämbetet, 1994), p.21.

Quality System

Origin of the review:	Function of the review	
	Supportive	Evaluative
Internal	QualW	
External	QualO	

William Massy and Andrea Wilger, NCPI Project 6 emergent conclusion

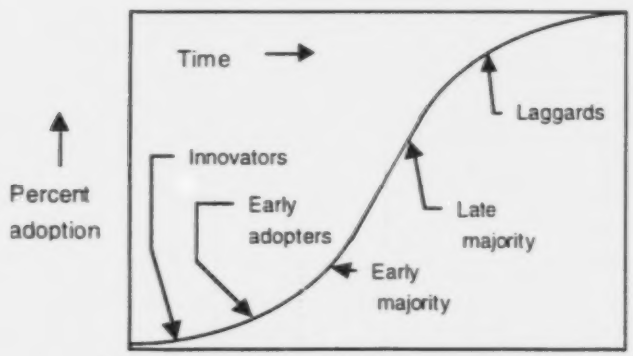
The right side of the Figure depicts our conclusion that a good quality system intersects all four quadrants of Trow's typology. The shaded oval represents the quality work (*QualW*) and quality oversight (*QualO*) elements of the quality system. The system is mainly internal and supportive, but it contains external and evaluative elements as well. Internal quality work is designed to produce improvement. However, quality shortfalls that are not mitigated in a reasonable period of time will elicit an evaluative response with appropriate consequences—hopefully within the department but at the decanal or institutional level if necessary. (Exemplary quality should produce positive consequences.) Bodies outside the institution's internal governance apparatus perform quality oversight. Such oversight should be mostly supportive but retain the capacity for evaluation and action.

The Adoption Cycle

Acceptance of the quality work concept and adoption of the requisite behavior don't take place overnight. Even under optimal conditions, it may take years to adapt the concept to a new setting, build expertise, and change the academic culture. People who seek a quick fix will be disappointed—all the more so until the prerequisites for quality work become firmly embedded. The adoption of innovations usually starts slowly and then accelerates once a critical mass of successful experience has been achieved. We should not expect the adoption of educational quality work to behave differently.

BEST COPY AVAILABLE

Figure 3. The Diffusion of Innovations



Adapted from Everett M. Rogers, *Diffusion of Innovations* (New York: Free Press, 1964), p. 61

Figure 3 presents the so-called "s-curve of adoption"—the classic diffusion curve for innovations—together with the category names usually associated with successive adopter groups. The *innovators*, which represent a few percent of the population of eventual adopters, are more likely to seek out and experiment with new ideas than those who adopt later. Usually they are part of informal information networks that include other innovators. The *early adopters*, the next 15 percent or so, may be moved to adopt once the innovators have perfected the innovation and demonstrated its benefits. Typically, they are more tightly connected to the local organization, where they often are viewed as opinion leaders. Members of the *early majority*, roughly the next third of the population, display less leadership than the early adopters but they are open to new ideas and tend to be respected by their peers. The *late majority*, the 33 percent who adopt after half the population has already done so, often are driven by social or economic pressures generated by the innovation rather than forward-looking vision. They are the followers, either through conservatism or because their attention was focused elsewhere during the earlier adoption stages. The last 15 percent or so, the *laggards*, resist adopting the innovation despite its advantages and the risk of becoming isolated from the population mainstream.

The adoption curve for quality work can be applied to a country's higher education system as a whole, to the departments and programs within a particular university, or to individual professors. Our research indicates that academic quality work has become well established at the innovation stage, but it appears that only a few countries, universities, and departments have progressed further.

Quality systems will always need an element of oversight, but in steady-state conditions such oversight will be of lesser importance than the quality work itself. But because quality work remains at an early stage of diffusion, quality oversight is especially important now. Effective oversight can stimulate the diffusion of quality work, although crude accountability exercises may well do the opposite. Some critics have urged that quality oversight should be deferred until local quality work has become better established. We believe this would be unwise. Well-designed quality oversight processes can safely precede the widespread adoption of quality work and they can accelerate its diffusion.

Quality Work and the Scholarship of Teaching

Ernest Boyer's *Scholarship Reconsidered* identified four distinct types of scholarship: the scholarship of inquiry (traditional research), the scholarship of integration, the scholarship of application (now known as the scholarship of engagement), and the scholarship of teaching. The book and its sequel, *Scholarship Assessed*, generated great interest but also some confusion about the scholarship of teaching. Recent work by Lee Shulman and his colleagues has clarified matters. Scholarship-based teaching is different from the scholarship of teaching. All scholarship, including the scholarship of teaching, must produce generalizable results that advance the field and can be reviewed by peers. Good teaching depends on all four kinds of scholarship, but it is not scholarship.

The same can be said about educational quality work. All kinds of scholarship can contribute to the five domains of quality work. The scholarship of teaching, with its emphasis on learning processes and outcomes measurement, is particularly relevant. The scholarship of teaching includes inquiry into quality work and how to improve it, but quality work itself is not scholarship.

The relation between quality work and scholarship has important implications for faculty reward and incentive systems. Generalizable and peer reviewed inquiry into quality work should be recognized as scholarship. More importantly, quality work itself should not be evaluated according to scholarly criteria. The criteria discussed herein should be used instead. Educational quality work falls under the rubric of teaching rather than scholarship. The value lies in its local contributions rather than its generalizability, and faculty should be rewarded accordingly.

Appendix B: Exemplary Quality Work in Denmark and Sweden

Andrea Wilger and I are compiling an inventory of exemplary quality work practices as part of our NCPI project on academic quality and productivity. By "exemplary" we mean practices that illustrate the evolving concept of quality work and which may be worthy of emulation. They are not necessarily "best practices," and while they do not occur frequently, we don't wish to suggest that they cannot be found at institutions other than the ones indicated. The descriptions are necessarily sketchy given the breadth of the subject and the resources at our disposal. More detailed benchmarking would be needed to verify the processes' details and efficacy in the particular context reported, but we are convinced that the descriptions are broadly correct and that these or similar processes can enhance the quality of education.

This Appendix presents exemplary practices that I identified during my visit to Denmark and Sweden in June 1999. The listed items represent "targets of opportunity" that came to light in my conversations and readings. Because my time was limited and I made no attempt to inventory all quality work practices, the list may exclude important work being done at one or another institution or department.

Copenhagen Business School

Quality is part of the strategic process at the CBS. For example, the Senate decided that by 1994 the School should have established:

- objectives and goals, means, and success criteria at all levels (faculties, departments, study programmes, and the CBS as a whole);
- methods for current evaluation of teaching, research, and administrative and managerial support functions; and
- an identification of major action areas in the short and long term, and relevant action plans.[73]

Internal program evaluation started in 1991, and all programs have been subject to evaluation since 1994. The relevant study committee discusses the design of the evaluation with the CBS Evaluation Unit, which conducts the evaluation as described below. Follow-up is the study committee's responsibility and is included in the committee's annual strategic report to the dean.

In 1994 the School engaged UK Professor Lee Harvey, a recognized higher education quality expert, to help develop a program of quality work—a program which they now are actively pursuing. Lee defines educational quality as including teaching and learning processes, fitness for purpose, and value for money. The program includes a shift from a teaching to a learning orientation, the systematic evaluation of teaching, and the use of stakeholder inputs in program design. During the last two years, Harvey's inputs and findings from stakeholder surveys have produced innovations that include a new emphasis on active learning and the development of transferable skills.

The academic senate has created a steering group to oversee and stimulate the quality work of staff and study boards. It is now understood that whereas the department retains the responsibility for research, the faculty-level study board is responsible for teaching and educational quality work.

A central evaluation unit, headed by Bente Kristensen, supports the study boards' evaluation functions. The unit works with the relevant study board and key teachers to develop and then process the student course evaluation questionnaires. A core set of questions applies to all courses, and additional questions may be targeted to specific courses. While the questions change from time to time, the key time series remain comparable as long at the course doesn't change significantly. In addition to course evaluations, the unit uses weekly diary studies (based on random samples so far as possible) to assess the number of hours that students devote to learning tasks. It stands ready to assist the study boards in surveying dropouts, graduates, and employers, and it conducts research on evaluation methodology.

The School has organized an educational development unit, which now employs six people. It provides assistance in all pedagogical methodologies including active learning and information technology. To compensate for the lack of pedagogical training in Ph.D. Programs, every new assistant professor must take a two-day general course that has been organized by the unit.

The Language Faculty has developed a separate educational development unit that addresses the unique culture and characteristics of teaching a second language. The unit's courses cover new research on the teaching of oral language skills, best practices in language teaching, use of information technology including the Web, and "more general pedagogical issues" like the cultural gaps and expectations gaps that can exist between students and teachers. Unit personnel have been successful in stimulating discourse on these subjects, but they believe additional staff development is needed: "In [teachers'] daily work it's hard to get real change."

The Language Faculty also has refined its student course evaluation process to make it more substantive and more responsive. The questionnaires include the usual tick boxes dealing with the program, the teaching, and the teacher, but they also ask students to write about themselves—their attendance, the amount of work they put into the course, and what they think about their own performance. Responsiveness has been improved by asking teachers to report on the evaluations at the next class, and then discuss any plans for change. The evaluation results (including the qualitative ones), and the teachers' summaries and responses, go to the course coordinator and eventually to the Study Board.

The Faculty of Economics has developed a "Principles of Pedagogy" document. The principles are grounded in research but oriented toward praxis. They focus on problem-oriented elements of learning—e.g., projects and cases where students need to organize their own learning process using the teacher as a guide. This represents more of a change in Denmark than in some other countries: for example, the Danish language didn't originally have a word for "learning" but rather allowed for "teaching" only.

The Faculty of Business Administration has developed a matrix of courses and skills, which maps the contributions of each course to writing, presentation, computing, problem-solving, and similar skills. The study board established a subcommittee consisting of five faculty members and five students to oversee this and other aspects of quality work. The Board is working to embed student time budgets within course design and raise student time on task from "25-37 hours" to 45 hours per week.

Danish Technical University

The university prides itself on training "practical engineers" who are well versed in problem solving as well as theory. It is developing new pedagogical methodologies and teaching materials to further this goal. Many of these innovations have relied on the application of research methods, and these present clear examples of the late Ernest Boyer's "scholarship of teaching."[74] The university's "Strategic Plan 1998" calls out the priority being given to such paradigm shifts.

The university's Center for Engineering Educational Development, headed by Palle Sorensen, has developed a position statement on general engineering competence:

> The engineer shall be capable of interpreting complex problem situations and of translating them into technical or non-technical solvable problems. He shall be able to draw up criteria for the selection of solutions, taking into consideration technical as well as non-technical facts and conditions.[75]

Implementing the statement will require that students better understand the central elements of the curriculum and that faculty better understand the students' personal qualifications and their ability to deal with complex problems.

The development of project-oriented courses and more open problem-solving approaches (multiple approaches are encouraged) provide examples of paradigm shift. Research on an improved evaluation and quality model represents another.[76] Both efforts seek to improve the linkages between theory and practice and students' ability to apply their knowledge in real-world situations.

The so-called "Laboratory Project" provides another example of how research methods are being applied to improve teaching and learning. A project-designed CD-ROM that describes and simulates an efficient industrial plant was given to a group of practicing engineers with different theoretical backgrounds. The respondents were asked not only to solve the problem, but also to describe how they used theory in finding the solution. The result: theoretical background did matter, and the insights gained from the experiment did provide important information for program design. The Laboratory Project CD also is used in the classroom to enhance active learning.

Göteborg University

Göteborg has been described as almost a federation of faculties. Hence, the faculty boards have full responsibility for academic quality and quality work. Initially the main task of the University's central quality unit, headed by Professor Berit Askling, was to work closely with the university's education research professors and assist the faculty boards. The approach was based on voluntary participation, drawn in part by the unit's strong research legitimacy. Once the national evaluation program was announced, the unit was charged with developing a quality assurance program. More recently it has broadened its mandate to include "competency development for academic teachers (in teaching and evaluations) [and] for academic leaders (heads of departments, study directors)."[77] The unit assumed responsibility for drafting Göteborg's self-evaluation report, which "gave it an opportunity to remind the faculties of the quality assurance programme."

Askling describes the unit's policy for quality work as follows:

- to create good preconditions for the operative staff [e.g., the faculty]
- to create arenas to be used for support and control [and to] to develop an infrastructure to be used for "horizontal learning" [i.e., benchmarking good practice across faculties and departments].

She describes the approach to horizontal learning as using the "very old tradition of the academic research seminar, which is not the same as measurement": get faculty and Ph.D. students to present what they're doing and then build up a database of good practice.

Appendix C describes the "System for Quality Assurance and Quality Development" put in place by Professor Bengt-Ove Boström and his colleagues in the Department of Political Science. This is the best such statement I've seen. It seeks to recognize and build on the faculty's professionalism in teaching and learning by channeling it into quality work. The system is a comprehensive one. It deals not only with teaching and learning as such, but also with administration and research. According to Boström:

> Making the teachers/scientists and the administrators realize that the quality of teaching and research in part relies on the quality of the administration is important. And it is also important that quality activities concerning teaching and research are in the same document, because the combination emphasizes that both are important parts of our mission.

His views on the purpose of academic leadership also are worth quoting:

- to provide the best possible circumstances for [academic] staff in doing their job (organization, training, premises, computers, administration, finance, etc.)

- to provide *arenas* for professional *stimulation* and *evaluation*, and to make sure these arenas are used (student evaluations, course conferences, "staff development discussions", publications, seminars, research conferences and other forms of scientific debate, etc.).[78]

Department members cannot demonstrate that quality work improves quality, but they believe that it does. They see quality work as a collective responsibility, and they have sufficient confidence to give it a priority similar to research. (The department has well-known researchers, but it has long been concerned about educational quality.) Perhaps the fact that Political Science at Göteborg competes for "commissioned education funds" from corporate sponsors contributes to this feeling. The department operates in the marketplace as well as in the academy, and good teaching pays off in tangible as well as intangible ways. It uses quality work as a vehicle for enhancing its position within and outside the University. Boström's good reputation with the faculty and the dean helps further the department's quality work. He is described as "managing the department in a good way," and he is trusted.

The Economics department's practice of paying a few students to work on course evaluation represents another exemplary practice. They make up the evaluation forms, tally the results, and (importantly) discuss the results with the teachers. This gets students involved in the process—which is not always the case elsewhere—and thus improves communication and buy-in.

Uppsala University

The University Board created a Committee on the Quality and Evaluation of Teaching in anticipation of audit. (Uppsala was among the first institutions to be audited.) The Committee, chaired by business Professor Lars Engwall, included representatives of the various faculties, two members of the student union, and two members to represent staff not affiliated with any specific faculty. Its main task was to assist the institutional leadership with work relating to quality and the evaluation of education. The University Board also decided that five percent of the grants given to the different faculties should be used for quality work, and the Committee consulted with the deans on the use of these funds. The Committee also established an international advisory board that included Berkeley (CA) Professor Martin Trow, a recognized quality expert.

Engwall describes Uppsala's basic concept of quality work as follows:

> Evaluating what comes out of [an education] system is not enough; it is equally important to analyse the conditions under which the outcomes are achieved and how the work is organised. Another important consideration is the need to evaluate results in a longer-term as well as a short-term perspective.[79]

Faculties and departments were expected to find the most appropriate organization and approach for their quality work.

The Committee's first initiative was to inventory the kinds of quality work going on in the various academic units. It found that many faculties had chosen to create special working groups. While all mirrored the central group to some extent, there were significant variations. For example, in the Language Faculty, the group:

> [h]as been entrusted with encouraging and supporting the departments in their quality work—for example, by arranging meetings of directors of undergraduate studies to discuss quality development and follow-up, taking Section-wide initiatives on teaching methods, and organising

> training seminars for the Section's teaching staff. The group also has a mandate to propose models for self-assessment and quality assurance, including methods for evaluating entire one-term courses and shorter modules, and to promote external validation of the Section's departments—for example, by means of peer reviews. In addition, it is expected to identify any problem areas and recommend suitable action.

The Faculty of Social Sciences' group "arranged a conference, for all the Faculty's directors of undergraduate studies, on teaching development, student recruitment, and the link between undergraduate education and the market."

The inventory also included departmental self-evaluations of course assessment methodology and, of particular importance, faculty-level reviews of departmental management processes.

> [The reviews] can be exemplified by the Faculty of Science and Technology's submission to the Quality Committee. The report suggested that "many departments have now become overgrown and are performing rather poorly, partly owing to inadequate management." Further, it expressed the view that:
>
> > Management at the departmental and faculty levels must be more subject to evaluation and more accountable that it has been up to now: "If staff lack confidence in their management, they will feel less enthusiastic and less inclined to make an extra effort." Furthermore, communication between managers and other staff needs to be improved. It is important to inform staff about what is being planned and what decisions already taken will mean in practice.

Engwall believes the committee was successful in initiating discourse about quality work at Uppsala. He attributes this to the committee's ability to get the attention of the deans. This was aided by strong faculty and student involvement, and by the institution head's strong support.

Stockholm University

Stockholm University presents another highly decentralized case. The University established a "Pedagogic Council" in 1989, well before audit, with the objective of "systematizing quality work." In 1991 it prepared a matrix laying out the kinds of quality work

being done by each department. The paper, which was construed as setting forth University goals for new forms of teaching and examination, was not well received. The established view held that most power belonged in the departments, with the central administration and even the faculties being relegated to supportive roles.

The central quality unit's director, Mona Bessman, recognized that no top-down strategy could work. Hence, she concentrated on "starting with the possible"—harvesting the low-hanging fruit. For example, the unit offers seminars that bring different approaches and theories about teaching and learning to the attention of faculty and students. These seminars seek to build intrinsic interest in quality work by competing with the less satisfactory elements of traditional disciplinary research. (One respondent characterized the work to be replaced as "mostly pointless except to create profits for a few Dutch capitalists.") Good quality work turns out to be academically meaningful and intellectually rewarding. Bessman describes her strategy as analogous to a microwave oven: "to heat from the inside."

Departments are expected to set their own goals for quality work. This is described as "both a strength and a weakness." One the positive side, it generates motivation and ownership and recognizes that only the department is close enough to the action to be effective. However, it requires a high degree of patience on the part of the central administration. Departments say they have always been working with quality, and the central administration didn't want to confront these objections even in the face of external audit. Instead, it wanted to work with the "good forces," to lead by example.

The external panel found many good things happening at the departmental level, but criticized the University for weak overall leadership. However, the University has stood by its position that it is essential to protect the departments' "free space" to develop their own agendas.

The central administration has not been passive, however. The Rector awards an annual "pedagogical prize" of SKR 20 thousand (about U.S. $2,400), which can be spent on work-related perquisites at the winner's discretion, to celebrate best practice as identified by students. In addition, the quality unit runs run a university-wide annual conference to share good practice. These programs, and the unit's ongoing support of the University's quality work leaders, are described as "highly effective" in promulgating good practice.

Perhaps most important, the administration asked all departments to comment constructively on the first-round external audit report and recommend actions to the Rector. The defense of free space coupled with dissemination of good practice appears to have paid off in unanimous departmental support for the following:

- informal involvement of students in the development and evaluation of teaching and learning processes

- the need for departments to crystallize their goals, to clarify and improve their policy documents.

In addition, there was substantial agreement on the need for departments to:

- do better at evaluating their teaching and learning programs and following up on the results

- get feedback from employees and former students.

Subsequently the University Board adopted the following program:

- affirmed decentralization

- departments to develop goals according to a common framework to be developed with broad consultation

- academic leadership to be proactive and meet regularly with departments on quality matters

- students to be engaged in quality work at the faculty and departmental levels

- evaluation processes to be improved and made more salient for students

- better information to be provided to students.

The course evaluation system has received particular attention. Policy requires that all courses be evaluated, but the process had become routine and sterile: students didn't take it seriously and the results were rarely acted upon. In one experiment aimed at making the system more meaningful, students who have just finished the course report their evaluations to new students and then the teacher describes what changes are being made.

The Board's initiatives have been organized according to the following timetable:

- 1998: Inventory and analyze quality work across departments, derive patterns, identify gaps. (This is reminiscent of the ill-fated 1991 "matrix," but the process now is on much firmer ground.)

- 1999: Conduct seminars and develop a self-assessment structure to help the departments see what to do.
- 2000: Departments and faculties will conduct their self-assessments.
- 2001: The National Agency will visit to Stockholm University for its second audit round.

These initiatives illustrate how repeated evaluation can function as a stimulus for learning even when the original ground was less than fertile.

Appendix C: "A System for Quality Assurance and Quality Development"

Department of Political Science, Göteborg University,
Professor Bengt-Ove Boström, Chair
(Bengt-Ove.Bostrom@pol.gu.se)

The Legitimacy of the System

The system for quality assurance and quality development was adopted by the Board of the Department of Political Science, May 26, 1998. This system is applicable to all Department activities unless in conflict with government statutes, decisions made by higher bodies within Gothenburg University, or subsequent Department Board decisions. It is a part of the quality assurance system of Gothenburg University.

Throughout this document, the word "colleague" refers to both formally employed staff members as well as to active Ph.D. students. The term "Department leadership" refers both to the Department Board and to the Head of Department. The relations between the Board and the Head of Department, as well as between other leadership positions, are specified in the documents which outline authority levels for Gothenburg University, the Social Science Faculty and the Department of Political Science respectively.

Goals, Results and "Modus Operandi"

The ultimate objective of the Department's quality work is to improve the results of our activities. In striving for this, quantitative goals must not be achieved at the expense of qualitative goals. Our quality work consists of three parts, in which we focus on goals, results, and "modus operandi" (i.e., our way of working) respectively.

In the first phase, our goals are deliberated and adopted. What do we want to achieve during a given period of time? The quality of such goals is assumed to be a function of the extent to which they have been deliberated. The demands on our Department increase as the ambition level of the goals rises.

In the second phase, the results of the Department are compared with adopted goals. To what extent are the goals achieved? In terms of quality, results improve as they near the adopted goals.

In the third phase, our "modus operandi" is evaluated. Are we working in an appropriate and reasonable manner, with respect to the goals? In terms of quality, our "modus operandi" improves when it becomes better suited to achieve our goals.

In sum, it is our goals and our "modus operandi"—and the regular undertaking of revision of these goals, results, and "modus operandi"—that will develop the quality of our results. This system for quality assurance and quality development specifies how such revision will be pursued.

Prerequisites

The "modus operandi" of the Department includes both the way in which individuals work and the organization and division of labor within the Department. This system considers both aspects, building on the following assumptions regarding prerequisites for positive development.

The possibilities for the Department's leadership to assist in achieving the Department's goals are improved if:

- the goals are clear, well-known and accepted by the colleagues of the Department;
- the leadership has control over the usage of resources on a continuous basis;
- the distribution of responsibilities and authority between leadership positions is clear and reasonable;
- the leadership is informed as to how colleagues experience their work, as well as the Department's goals and "modus operandi"; and
- the leadership is aware of the qualifications and ambitions of their colleagues.

As individual colleagues, our possibilities to contribute to the quality development of the Department are enhanced if:

- we are familiar with the organization, the activities, and the goals of the Department, as well as our own roles in the organization and in carrying-out its activities;
- our responsibilities and authorities are well-defined;
- we have the proper competence and the proper equipment to perform our tasks;

- we perceive our tasks as meaningful; and
- our tasks are reasonable with respect to working hours.

This system for quality assurance and development shall help to fulfil these prerequisites. Also, it shall contribute to the accomplishment of important tasks and help to insure that the Department uses its resources, both human and material, in an efficient way. The latter involves the efficient usage of relevant technological equipment and other tools, as well as of our facilities. Finally, this document outlines procedures for the regular revision of the system itself.

Revision

The revision work involves evaluating our goals, results and "modus operandi," and implementing those changes/measures that are deemed appropriate. This shall be conducted with such methods and in such a spirit that the development of colleagues and the Department is best promoted.

The following aspects shall be subject to regular revision:

- general goals and objectives
- organization and utilization of resources
- budget
- the relation of colleagues to the Department
- research and the Ph.D. program
- undergraduate education
- external education (i.e., teaching under external contract)
- external information
- administrative functions

Goals and Objectives

Every third year, the Department Board shall reconsider the general goals of the Department, as well as evaluate results with respect to the currently valid decisions regarding goals and objectives. On these occasions, the Board shall also evaluate and reconsider the quality assurance system of the Department. The revision shall be preceded by a meeting between the Head of the Department, Directors of Studies, Ombudsman for Gender Equality, Student Representative, as well as the Department's professors, and it can also be preceded by an expanded deliberation process. It is the responsibility of the Head of the Department to gather information on which the review process is based. It is the responsibility of the Board to assure that the Department is working in accordance with the goals and objectives, and the quality assurance system.

The gender equality goals of the Department are established in a special gender equality program, which shall be the subject of revision every third year.

The educational goals for courses and modules are specified in special documents established by sub-faculty, and faculty boards, on the initiative of the Department Board. The development of these plans is conducted by lecturers on a continuous basis. The teaching staff for the Ph.D. program, and for undergraduate education, shall each collectively revise their course plans annually (see below).

Organization and Utilization of Resources

Every third year, the Head of the Department shall make an evaluation of the organization of the Department with respect to: (1) the distribution of tasks, information, responsibilities and authorities, and (2) how material resources and human resources are being used. This evaluation shall also include a review of the delegation of authority. The evaluation shall be reported to the Department Board. It would be most appropriate to conduct this evaluation during the semester immediately prior to the evaluation of goals and objectives—it can thus be part of the information on which such decisions are based.

Furthermore, the availability and distribution of office space, computers and other kinds of equipment—as well as the need for human resource development—shall be revised at the beginning of each financial year.

The Head of the Department and the Board shall take appropriate steps on the basis of these completed evaluations.

Budget

Before each financial year, the Board shall decide on a budget for the Department. If necessary, it shall also make major corrections of the budget in its ongoing work throughout the year. Before the annual budget decision, the Board shall have access to information concerning the financial results of the preceding year.

Relations between Colleagues and the Department

In order to improve the possibilities for individual colleagues to develop professionally, and in order for the Department leadership to collaborate with colleagues to accomplish Department goals, special "staff development" discussions shall be arranged for all colleagues and Ph.D. students. It is the responsibility of the Head of the Department to ensure that such discussions are held annually for Ph.D. students, and every second year for other colleagues. Guidelines for these conversations can be found in a special document.

Research and the Ph.D. Program

The quality of research conducted at the Department is assured by means of collegial scrutiny prior to publishing, in reviewed articles, in reviews conducted during the decision process for promotions, in doctoral disputations, in seminars and in other modes of scientific debate. To a large extent, scientific quality control is located outside the internal activities of the Department.

The Department shall strive to put all its scientific research products under internal as well as external collegial scrutiny. The Department shall set up a database consisting of all scientific reports produced at the Department from July 1, 1998, and onwards. The database shall contain information as to in which contexts reports have been the subjects of scrutiny, as well as information about expert scrutiny tasks externally assigned to colleagues (appointment processes, research grants, doctoral and licentiate dissertations, or roles as experts on government commissions).

According to Swedish law, it is part of the universities' research mission to "spread information about teaching and research and how the knowledge and experience generated at the universities can be applied." In order to create a basis for evaluation of how the Department fulfils this task, the database shall include information about colleagues' involvement in such information endeavors.

It is the responsibility of individual colleagues to provide the person in charge of the database with the relevant data. The database shall be the basis of the annual report of

the Department. In combination, the database and the annual report form the foundation for an overall evaluation of research activities prior to the Board's annual decisions regarding goals and objectives.

Each year, the general and specialized research seminars shall be evaluated by the respective participants. These evaluations may be done in writing or by means of discussion during regular seminars or in special planning seminars. All courses in the Ph.D. program shall be evaluated by the students each time courses are held. The Director of Studies and the affected teachers shall be informed about the contents of the evaluations. Evaluations done in writing must be made available to the students. The entire course package of the Ph.D. program shall be evaluated by the Director of Studies and those students that have recently finished the course program.

The dissertation work and the advising of the Ph.D. students shall be planned and evaluated by both the adviser and the student in conjunction with the "staff development" discussions (see above). The discussions shall precede the advisers' annual report to the Director of Studies concerning the academic progress of the student. In addition, the discussions shall include the future post-doctoral situation of the student. The overall advising situation, the financial situation of the Ph.D. students, as well as the Ph.D. program in general, are to be evaluated annually by the Board of Advisers.

At least every third year, the Department shall establish contact with another university department with the purpose of gaining knowledge that can benefit our Ph.D. program. Such contacts can involve "benchmarking," external revision, and mutual exchange of experiences.

Undergraduate Education

Similar to the Ph.D. program, undergraduate education shall be evaluated on a regular basis through internal and external activities.

Each time they are held, undergraduate modules shall be evaluated by the students. The Director of Studies and the participating teachers shall be informed about evaluation outcomes. Evaluations done in writing are to be made available to the students.

Each module will be the subject of a development conference at least once every third year. In addition to the participating teachers, two colleges and two student representatives shall participate in such conferences. Conference practicalities are specified in a special document.

For each module, there shall be a "log book," containing previous and current documents specifying its goals, its teaching plan as well as protocols from earlier development conferences. The existence of such a folder makes the changing of teachers easier, and it serves as an information resource for future development conferences.

Undergraduate courses, equipment, and facilities are to be evaluated by student representatives and the Director of Studies at the end of each semester.

Before each academic year, the Department shall arrange an undergraduate education conference where all modules are presented by the participating teachers, and where the individual modules are discussed in relation to one another as well as in relation to the Ph.D. program. As a basis for discussions, course plans, literature lists, and statistics pertaining to the flow of students etc., shall be available.

Every third year, undergraduate education shall be evaluated with respect to its relevance for the students' future labor market situation. Such evaluations can take on different profiles on different occasions (for instance, survey investigations of former students, or potential employers).

At least every third year, the Department shall establish contact with another university department, with the purpose of gaining knowledge that can benefit our undergraduate education. For example, such contacts can involve "benchmarking," external revision, and mutual exchange of experiences.

External Education

When the Department agrees to carry out education on behalf of another principal, the affected teachers and Director of Studies at the Department of Political Science shall discuss the results of conducted evaluations with the principal.

External Information

Every autumn semester, the Board shall consider and decide the extent to which the external information of the Department needs to be improved for the coming academic year.

Administrative Functions

Before each academic year, the Head of the Department and the administrative staff shall evaluate the past year, and plan the coming year with respect to administrative work. Representatives of both teachers/researchers and students shall participate in the

evaluation. The Board shall appoint these representatives during its last spring semester meeting. Moreover, representatives of the Administration from the Faculty of Social Sciences and the Office of Education shall be invited and they shall be given the opportunity to report on activities for the coming year. In addition, this allows for the exchange of viewpoints concerning contacts between the Department and the Faculty Administration/Office of Education.

Finally, before the end of each semester, the undergraduate students of the Department shall be given an opportunity to evaluate in writing those administrative functions with which they have come in contact. It shall be carried out at the same time as the evaluation of courses, and it is to be used as background information in that process.

Affected employees and supervisors shall discuss evaluation results. To the extent that results affect the Head of the Department, they shall be dealt with by the Board.

Immediate Measures

The periodical revision processes outlined in this document must never hinder immediate problem solving, which is deemed beneficial to the Department. Responsibility lies with each individual colleague concerned to carry this through.

Endnotes

1. Jette Kirstein, "Trends in Learning Structures in Higher Education" (Confederation of European Rectors Conferences, July 1999), Section 4.
2. The entities' primary representatives were as follows (in the order visited): Danish Centre for Quality Assurance and Evaluation of Higher Education, Copenhagen (Christian Thune, Director: Christian.Thune@eval.dk); Copenhagen Business School (Bente Kristensen, Vice President in charge of quality programs: Bk/ledsek@cbs.dk); Danish Technical University, Copenhagen (Palle Sorensen, director of the educational development unit: Ps@cdm.dtu.dk); Göteborg University, Sweden (Berit Askling, Professor of Education and director of the central quality unit: Berit.Askling@ped.gu.se); Uppsala University (Lars Engwall, Professor of Business Studies and founding chair of the Quality Committee: Lars.Engwall@fek.uu.se); Swedish Parliament (Per Unckel, MP and former Minister of Education: Per.unckel@riksdagen.se); National Agency for Higher Education, Stockholm (Staffan Whalén, Coordinator of the Quality Audit Programme: Staffan.Wahlen@hsv.se); Stockholm University (Mona Bessman, Educational Consultant, Educational Development Unit: Mona.Bessman@pu.su.se). I also met with Stig Hagstrom (former Chancellor of the Swedish higher education system: hagstrom@stanford.edu) at Stanford University.
3. Staffan Wahlén (1998), "Is there a Scandinavian model of evaluation of higher education?" (*Higher Education Management*), p. 28. The paper argues for commonality not only between Sweden and Denmark, but among all four Scandinavian countries.
4. *Ibid.*, p. 28. Emphasis in the original.
5. Berit Askling (1998a), "Institutional government and management in a context of policy and system changes—the Swedish case" (Göteborg University, Department of Education and Educational Research, Report 1998:02), p. 17.
6. *Ibid.*
7. This is consistent with the points made in Policy Perspectives, "A Transatlantic Dialog" (University of Pennsylvania, Institute for Research on Higher Education, Pew Higher Education Research Program, Vol. 5, 1993).
8. Wahlén (1998), p. 35.
9. Hagstrom returned to Stanford in early 1999 at the completion of his term. His director of quality evaluation, Wahlén, remains in place. Per Unckle remains in Parliament as a minority member and chair of the Committee for Constitutional Reform.

10. Per Unckle quoting Stig Hagstrom.

11. Wahlén (1998), pp. 370-71.

12. Hagstrom interview.

13. Large-scale subject-level assessments were being conducted by the Higher Education Funding Council of England (HEFCE). These assessments had a distinctly summative flavor and were in fact intended to drive funding.

14. E.g., Staffan Wahlén, Coordinator of the Quality Audit Programme, has been active in the higher education quality profession for most of the decade.

15. The Hong Kong University Grants Committee (UGC) uses the term "Teaching and Learning Quality Process Review (TLQPR)," though in fact the reviews are quality process audits as the term is used here. See William F. Massy, "Teaching Learning Quality Process Review: The Hong Kong Programme" (*Quality in Higher Education*, 3[3], 1997), pp. 249-262; William F., Massy & Nigel J. French, "Teaching and Learning Quality Process Review: A Review of the Hong Kong Programme" (presented at the fourth conference of the International Network for Quality Assurance Agencies in Higher Education: INQAAHE, South Africa [May 1997]), William F. Massy and Nigel J. French, "Teaching and Learning Quality Process Review: What Has the Programme Achieved in Hong Kong?" (presented at the 1999 conference of the International Network for Quality Assurance Agencies in Higher Education – Santiago, Chile, 2-5 May 1999). I serve on the UGC and chaired the TLQPR steering committee and site visit teams.

16. Malin Östling, "Experiences from the Evaluation Centre" (Swedish National Agency for Higher Education internal paper, Dec-Jan 1997-98), footnote 1.

17. Christian Thune, "The Alliance of Accountability and Improvement: the Danish experience" (*Quality in Higher Education*, Vol. 2, No. 1, 1996), p. 23.

18. Thune also served as international expert for Hong Kong's quality process audits.

19. Thune (1996), p. 24, emphasis added.

20. *Ibid.*, p. 24.

21. *Ibid.*, p. 22.

22. *Ibid.*, p. 22.

23. "The National Quality Audit of Higher Education in Sweden" (Stockholm: National Agency for Higher Education, 1996), p. 8.

24. C.f., Christian Thune, "The European Systems of Quality Assurance: Dimensions of Harmonisation and Differention" (paper presented at the International Dissemination Conference on Quality Assessment in Higher Education [OECD/IMHE], Mexico City, December 1-2, 1997), p.13.

25. For more complete accounts see the many publications and papers cited herein, especially "The National Quality Audit of Higher Education in Sweden" (1996), Thune (1996), Wahlén (1998), Östling (1997-98), and the agencies Web sites.

26. Östling (1997-98), p. 3.

27. Thune (1996), p. 27.

28. Wahlén (1998), p. 31.

29. "The National Quality Audit of Higher Education in Sweden" (1996), p. 29.

30. *Ibid.*, p. 29.

31. Malin Östling, "Self-Evaluation for Audit—A Tool for Institutional Improvement?" (Swedish National Agency for Higher Education, Department of Evaluation and Quality Audit; presented at the 19th EAIR Forum, Warwick, 1997), p. 15.

32. *Ibid.*, p. 5.

33. Both sets of questions are listed in "The National Quality Audit of Higher Education in Sweden" (1996), at p. 30.

34. Östling (1997), p. 9. Only a few of both countries' reports are available in English.

35. Dorte Kristoffersen and Christian Thune (1999), "Guarding the Guardian: The Evaluation of the Danish Centre for Quality Assurance and Evaluation of Higher" (paper presented at The Fifth Biennial Conference of the International Network for Quality Assurance in Higher Education, May, Santiago, Chile), p. 4.

36. Östling (1997), pp. 11-12. Not all the respondents' quotes have been included and the grammar has been smoothed in a few cases to mitigate shortfalls in translation.

37. Bjørn Stensaker, "External Quality Auditing in Sweden: Are Departments Affected?" (Oslo: Norwegian Institute for Studies in Research and Higher Education, study commissioned by the Swedish National Agency for Higher Education; forthcoming in *Higher Education Quarterly*, 1999a), pp. 15-16. Emphasis in the original. The quotation within the text refers to John Brennan, Mark Frederiks, and Tarla Shah, "Improving the Quality of Education: The Impact of Quality Assessment on Institutions" (Milton Keynes, Quality Support Centre and the Higher Education Funding Council for England, 1997), p. 12.

38. Östling (1997), p. 13.

39. Östling (1997), p. 8.

40. The reports I reviewed were as follows. Denmark: The Political Science Study Programme and Public-administration and Management Study Programmes; The Mathematics, Physics and Chemistry Programmes. Sweden: University of Uppsala; Stockholm University; Mid-Sweden University College. English versions of the

Danish reports can be found at the Evaluation Centre web site. Only a few of the Swedish reports have been translated to English and I obtained them in hard copy.

41. Östling (1997), p. 8.
42. *Ibid.*, p. 11.
43. *Ibid.*, pp. 9-10.
44. Östling (1997-98), p. 14.
45. *Ibid.*, p. 15.
46. Östling (1997), p. 9.
47. *Ibid.*, p. 9.
48. Stensaker, personal communication.
49. *Ibid.*, p. 19.
50. *Ibid.*, p. 19.
51. Stensaker reports that his forthcoming article in *Quality in Higher Education* includes a similar conclusion.
52. Stensaker (1999a), p. 13.
53. Stensaker suggested this in a personal communication, and I agree.
54. Östling (1997-98), p. 16.
55. Ibid., pp. 19-20.
56. Stensaker (1999a), p. 14.
57. *Ibid.*, p. 15.
58. Wahlén (1998), p. 38.
59. *Ibid.*, p. 38.
60. Bjørn Stensaker, "Quality as Discourse: An analysis of External Audit Reports in Sweden 1995-1998" (Oslo: Norwegian Institute for Studies in Research; paper presented at the 21st EAIR Forum, Lund University, August 1999b), p. 4.
61. *Ibid.*, p. 6.
62. *Ibid.*, p. 5.
63. *Ibid.*, pp. 6-7. My own reading of the reports support this view.
64. Berit Askling (1998b), "Quality Work in Swedish Universities in a Period of Transition" (Göteborg University, Dept. of Education and Educational Research, Report No. 1998:04; paper presented at the Finnish Higher Education Council's Seminar, "Quality Work in Universities," October 20, 1998), p. 12.

65. John Brennan, "Authority, Legitimacy, and Change: The Rise of Quality Assessment in Higher Education" (paper presented at the CERI/IMHE Conference in Paris; December, 1995). Brennan was chair of the UK 's Academic Quality Council.

66. Askling (1998a), p. 76.

67. Ibid., p. 76.

68. Askling (1998b), p. 16.

69. Stensaker (1999a), p. 23.

70. Christian Thune was appointed Director of the Institute on September 12, 1999.

71. Burton Clark, Creating Entrepreneurial Universities: Organizational Pathways of Transformation (New York, Pergamon, 1998), p. 148; quoted by Askling (1998b), p. 15.

72. Frans van Vught (1999), "Quality Audits and the Chambers of the Academic Heart" (presentation to the University of North Carolina's Conference on Academic Audit, June). Van Vught was founding director of the Center for Higher Education Policy Studies at the University of Twente (NL) and currently serves as Rector of the University.

73. Bente Kristensen, "Increased workload or synergy in evaluations and audits: Quality management at the Copenhagen Business School" (paper presented at the 18th Annual EAIR Forum, Budapest, Hungary, August 25-28, 1996), p. 5.

74. Ernest Boyer, *Scholarship Reconsidered* (Princeton, NJ: Carnegie Foundation for the Advancement of Teaching, 1991).

75. "Proposal Center for Engineering Education Development with a definition of a new, general engineering competence as a basis for DTU's paradigm shift in the educational programmes and teaching", Technical University of Denmark (Copenhagen: undated), p. 2.

76. Camilla Rump, Arne Jakobsen, and Torkil Clemmensen, "Improving Conceptual Understanding Using Qualitative Tests" (Center for Engineering Education Development, Technical University of Denmark, November 1998).

77. This and the subsequent quotations are from Askling (1998b), p. 9.

78. Bengt-Ove Boström (1999), presentation slide.

79. Lars Engwall, "Institutional Responses to Quality Assessment: A Swedish Approach to Quality in Education, the Case of Uppsala University" (Paris: OECD, Organisation for Economic Cooperation and Development, Centre for Educational Research and Innovation, Programme on Institutional Management in Higher Education, CERI/IMHE/CG[95]20, 4-6 December 1995), p. 11. The next two quotations are from pages 12 and 13 of the same document.